

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 14:55:01 ; Search time 207 Seconds

(without alignments)
4942.493 Million cell updates/sec

Title: US-09-975-253-1

Perfect score: 2699

Sequence: 1 ggttcagctgccgcagcgc.....aaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_n2p.model -DEV=xl
-O=/cgn2_1/USPTO_spool_p/US09975253/runat_28102004_111509_19617/app_query.fasta_1.1607
-DB=A_Geneseq_238Sep04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09975253 @CGN_1_1_227 @runat_28102004_111509_19617 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_238Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	86.2	427	3 AAY87783	Aay87783 Human IRF
2	2326	86.2	427	4 AAM78986	Aam78986 Human pro
3	2326	86.2	427	6 ABG76085	Abg76085 Human int
4	2326	86.2	427	7 ABW02028	Abw02028 Human int
5	2326	86.2	427	7 ADG47651	Adg47651 Human IRF
6	2304	85.4	427	2 AAY15102	Aay15102 Modified
7	2299	85.2	455	7 ADG77016	Adg77016 Human nuc
8	2018	74.8	392	4 ABB11963	Abb11963 Human IFN
9	2018	74.8	392	4 AAM79970	Aam79970 Human pro
10	1766.5	65.5	542	2 AAY15104	Aay15104 Chimeric

11	546	20.2	102	3 AAG00953	Aag00953 Human sec
12	504.5	18.7	488	7 ADD01173	Add01173 Human nuc
13	456.5	16.9	516	8 ADJ75465	Adj75465 Marker ge
14	456.5	16.9	516	8 ADJ75413	Adj75413 Marker ge
15	456.5	16.9	516	8 ADL83070	Adl83070 Human PRO
16	456.5	16.9	516	8 ADN04327	Adn04327 Antipsori
17	456	16.9	844	8 ADP29911	Adp29911 Human sec
18	454	16.8	503	2 AAY15103	Aay15103 Modified
19	448	16.6	503	7 ADES8665	Ades8665 Human PRO
20	448	16.6	503	7 ADES8669	Ades8669 Human PRO
21	448	16.6	503	8 ADJ75462	Adj75462 Marker ge
22	448	16.6	503	8 ADJ75410	Adj75410 Marker ge
23	446	16.5	503	4 AAE09329	Aae09329 Human int
24	444.5	16.5	474	8 ADJ75411	Adj75411 Marker ge
25	444.5	16.5	474	8 ADJ75463	Adj75463 Marker ge
26	436	16.2	457	7 ADES8667	Ades8667 Rat Prote
27	436	16.2	457	7 ADES8667	Ades8667 Rat Prote
28	436	16.2	457	8 ADJ76202	Adj76202 Marker ge
29	436	16.2	457	8 ADJ76237	Adj76237 Marker ge
30	399.5	14.8	94	5 ABB57389	Abb57389 Rat mucoc
31	399	14.8	77	4 AAM24435	Aam24435 Human EST
32	397	14.7	450	2 AAR99427	Aar99427 Human Lym
33	397	14.7	450	2 ABU03802	Abu03802 Human exp
34	397	14.7	450	6 ABU03798	Abu03798 Human exp
35	397	14.7	450	8 ADQ18306	Adq18306 Human sof
36	390.5	14.5	451	6 ABU03805	Abu03805 Human exp
37	390.5	14.5	451	6 ABU03799	Abu03799 Human exp
38	390.5	14.5	451	8 ADM67227	Adm67227 Murine ad
39	389.5	14.4	449	8 ADM67226	Adm67226 Murine ad
40	389.5	14.4	451	2 AAM38426	Aam38426 Human mul
41	389.5	14.4	451	6 ABU03804	Abu03804 Human exp
42	389.5	14.4	451	6 ABU03800	Abu03800 Human exp
43	389.5	14.4	451	6 ABU03807	Abu03807 Human exp
44	389.5	14.4	451	6 ABU03803	Abu03803 Human exp
45	389.5	14.4	451	8 ADL83104	Adl83104 Human PRO

ALIGNMENTS

RESULT 1

AAY87783

ID AAY87783 standard; protein; 427 AA.

XX

AC AAY87783;

XX 24-AUG-2000 (first entry)

DT Human IRF3 protein.

DE Human; ADA2; cytostatic; gene therapy; treatment; cancer; IRF3.

XX Homo sapiens.

OS Key

PH Location/Qualifiers

FT Misc-difference 8

FT /label= unidentified

XX US6054289-A.

XX 25-APR-2000.

XX 30-AUG-1996; 96US-00705771.

XX 30-AUG-1995; 95US-0002993P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA;

XX WPI; 2000-338491/29.

XX N-PSDB; AAA39474.

XX New polynucleotide encoding human AD2 is useful for treating cancer and

for isolating cDNAs and genes having similar biological activity.

Disclosure; Col 59-62; 54pp; English.

This invention describes a novel polynucleotide (1) encoding human ADA2. The products of the invention have cytosstatic activity and can be used for gene therapy. (1) is useful for treating cancer; as primers and probes for isolating full length cDNA and genes having similar biological activity. This sequence represents the human IRF3 protein described in the method of the invention

Sequence 427 AA;

Document Scores:			
No.:	1.9e-166	Length:	427
s:	2326.00	Matches:	427
Similarity:	100.00%	Conservative:	0
Local Similarity:	100.00%	Mismatches:	0
/ Match:	68.18%	Indels:	0
	3	Gaps:	0

9-975-253-1 (1-1426) x AAY87783 (1-427)

[illegible]

Db	241	TtpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr	260
QY	827	GTGAGCGATGTGCTCAGCTGCTGGTGGGGGACTGGCTCTCTCGCGCGCGGACGTGG	886
Db	261	ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTyrPArgIaGlyGlnTyr	280
QY	887	CTCTGGGCCCCGCGCTGGGCACTGCCACATACCTGGGCACTGAGCGAGGAGCTGCTC	946
Db	281	LeuTyrAlaGlnArgLeuGlyHisCysHisThrTyrTyrPalaValSerGluGluLeuLeu	300
QY	947	CCCAACACGGCGCATGGCGCTGTATGGCGAGTGCCCCCAAGGACAAGGAAGAGGCGTGT	1006
Db	301	ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe	320
QY	1007	GACCTGGGGCCCTTCATTGTATGACTCATACCTTACGTCGGNAGGAAGCGACGCTCACCA	1066
Db	321	AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro	340
QY	1067	CGCTATGCGCTCTGCTGTTCTGTGTGGGGAGTCAATGGCCCGAGGACCAAGCCGTGGACCAAG	1126
Db	341	ArgTyrAlaLeuTyrPheCysValGlyGluSerTyrProGlnAspGlnProTyrThrLys	360
QY	1127	AGGCTCGTATGTCATAGTTGTGCCCACTGGTGCCTCAGGCGCTTGGTAGAAATGGCCCGG	1186
Db	361	ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg	380
QY	1187	GTAGGGGTGCTCTCCCTCGGAGAACTGTGGACCTGCACATTTCCACAGCCACCCCA	1246
Db	381	ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro	400
QY	1247	CTCTCCCTCACCTCCGACCAGTACAGGCGCTACTGCAGGACTTGGTGGAGGCGATGGAT	1306
Db	401	LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp	420
QY	1307	TTCAGGGCCCTGGGGAGAGC	1327
Db	421	PheGlnGlyProGlyGluSer	427
RESULT 2			
AAW78986			
ID	AAW78986	standard; protein; 427 AA.	
XX	AAW78986;		
AC			
XX			
DT	06-NOV-2001	(first entry)	
XX			
DE	Human protein SEQ ID NO 1648.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation.		
XX			
OS	Homo sapiens.		
XX			
PN	WC200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US004098.		
XX			
PR	03-FEB-2000; 2000US-00496914.		
PR	27-APR-2000; 2000US-00560875.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y.		

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX
 DR N-PSDB; AAKS2119.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 3986; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 427 AA;

Alignment Scores:
 Pred. No.: 1.9e-166 Length: 427
 Score: 2326.00 Matches: 426
 Percent Similarity: 99.77% Conservative: 0
 Best Local Similarity: 99.77% Mismatches: 1
 Query Match: 96.18% Indels: 0
 DB: 4 Gaps: 0

US-09-975-253-1 (1-1426) x AAK78986 (1-427)

QY 47 ATGGGAACCCCAAGCAGCGNCTCTGCGCTGGCTGGTGTGCGAGTGCAGCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGln 20
 QY 107 CTGGAGGGCGTGGCTGGTGAACAGAGCCGCGCTTCCGCTCCCTTGGAGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACGCGAGATCGACAGCAGGAGATTTCGGAATCTTCCAGCGCTGGGCGGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGTGTGCATATGTCCTGGGAGGAGGATAAGCAGACCTGCAACCTGGAGAGGAATTC 286
 Db 61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGGTTCCGTTTACGAGACCGGAGCAGGACCCCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CAGCAGCCACATAAATCTACGAGTTGTGCACTCAGGAGTGGGACTTTTCCAGCCA 406
 Db 101 HisAspProHisLysIleTrpGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCCGACACCAATGTGGAGGAGTACTTCTGATACCCAGGAAGACATCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGAGTTACTGGTAAATAGTGTTCGCCCACTCCAGATCCGGAACCCCGAGCGTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGGGAGCCCGCCAGCTTGGACAATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180

QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCCACTGAAGCGGCTGTGTGGCCGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTCGAGGTGACAGCTTCTACCGGGGCGCCCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTrpArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGCGCTGGCGCTGGTGGCTCCGAGTGGGAGACAGAGCGCTCCCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCCAGACCCCTGGCATGTCCCTGACACAGCGGAGTGATGAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGATGTGTGAGTCTGCTGGTGGGAGTGTCTCTGGCGGGCGGCGGACATGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCGCCAGCGCTGGGCGCTGCCACACATCTGGGCGAGTGCAGCAGGAGTCTCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTrpAlaValSerGluGluLeuLeu 300
 QY 947 CCCAACAGCGGCGCATGGCGCTGTGCGGAGTCCCAAGGACAAGAGGAGGCGGTGTTT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTCAGGAAGGAGGCGGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCGCTCTGCTTGTGGGAGTGTGCTGGGCGGAGTGTGCGGCGGAGGAGGAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTCGTGATGCTCAAGTGTGTGCCACGCTGCTCAGGCGCTTGGTGAATATGCGCCGG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGCTCTCTCCCTGGAGAAATCTGTGGACCTGTCACATTTCCACAGCCACCA 1246
 Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCTCCGACCGAGTACAGGCTACCTGCGAGGACTTGGTGGAGGCGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTrpLysAlaTrpLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGGCGCTGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427
 XX
 RESULT 3
 ABG76085
 ID ABG76085 standard; protein; 427 AA.
 XX
 AC ABG76085;
 XX
 DT 09-MAY-2003 (first entry)
 XX
 DE Human interferon regulatory factor 3, IRF3.
 XX
 KW Human; interferon regulatory factor 3; IRF3; gene therapy; HIV infection;
 KW transcription protein; viral infection; Streptococcus pyogenes;
 KW human T-cell lymphotropic virus type 1; Epstein-Barr virus; scabies;
 KW encephalitis; yellow fever; bacterial infection; Neisseria gonorrhoea;
 KW gingivitis; syphilis; parasitic infection; Amebiasis; Babesiosis;
 KW malaria; deficient serum immunoglobulin production; recurrent infection;
 KW immune system dysfunction; joint infection; bone infection; malignancy;
 KW skin infection; parotid gland infection; blood-borne infection; epilepsy;
 KW autoimmune disease; inflammatory disorder; primary immunodeficiency;
 KW vaccine adjuvant.

OS Homo sapiens.

XX Key Location/Qualifiers

PH Misc-difference 8

FT /label= Val, Ile, Leu, Phe

FT /note= "Encoded by NTC"

FT 141.147

FT Region

FT /label= Nuclear_export_signal

FT 198.381

FT Domain

FT /label= Interferon_regulatory_factor_association_domain

FT 382.407

FT Region

FT /label= Phosphorylation_region

FT 408.427

FT Domain

FT /label= Autoinhibitory_domain

XX US2002164694-A1.

PN 07-NOV-2002.

XX 12-OCT-2001; 2001US-00975253.

XX 30-AUG-1995; 95US-0002993P.

PR 30-AUG-1996; 96US-00705771.

PR 13-OCT-2000; 2000US-0239936P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Pith-Rowe P;

PI WPI; 2003-298697/29.

DR N-PSDB; ABX12065.

XX New polynucleotide, useful for preventing, treating or ameliorating HIV disease.

PT Claim 1; Fig 1; 82pp; English.

XX The invention relates to an isolated human interferon regulatory factor 3 IRF3 polynucleotide. The interferon regulatory factors are transcription proteins. The polynucleotide or polypeptide is useful for preventing, treating or ameliorating HIV disease. The polynucleotide or polypeptide is also useful for treating viral infections e.g. human T-cell lymphotropic virus type I and Epstein-Barr virus and related diseases e.g. encephalitis and yellow fever; bacterial infections e.g. Streptococcus pyogenes and Neisseria gonorrhoea and related diseases e.g. gingivitis and syphilis; parasitic infections e.g. Amebiasis and Babesiosis and related diseases e.g. scabies and malaria; deficient serum immunoglobulin production; recurrent infections; immune system dysfunction; infections of the joints, bones, skin and/or parotid glands; blood-borne infections; autoimmune diseases; inflammatory disorders; malignancies; primary immunodeficiencies; epilepsy. The polypeptide is also useful as a vaccine adjuvant. The present sequence represents the amino acid sequence of the human interferon regulatory factor 3, IRF3

XX Sequence 427 AA;

SQ

Alignment Scores:

Pred. No.:	1.9e-166	Length:	427
Score:	2326.00	Matches:	427
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	86.18%	Indels:	0
DB:	6	Gaps:	0

US-09-975-253-1 (1-1426) x ABG76085 (1-427)

QY 47 ATGGGAACCCCAAGCCACCGNCTGCTGCTGGTGTGTCCAGCTGGACCTGGGGCAA 106

DB 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20

QY 107 CTGGAGGCGCTGGCTGCTCAACAGAGCCGACCGCTCCGATCCCTTGAAGCAC 166

DB 21 LeuGlyGlyAlaIleAlaIleValAsnLysSerArgThrArgPheArgLeuProTrpLysHis 40

QY 167 GGCTACGGCAGGATGCACAGCAGGAGGATTTTCGGAATCTTCCAGGCTGGCGCCAGGCC 226

DB 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60

QY 227 ACTGGTGCATATGTTCCCGGAGGAGTAAGCCAGACTGCTCCAACTCTGGAAGAAATTC 286

DB 61 ThrGlyAlaIleValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80

QY 287 CGCTCTGCCTCAACCGCAAGAGGTTGCGTTTACGAGAGGACCGGACAGGACCCCT 346

DB 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100

QY 347 CAGACCCACATAAATCTACGAGTTGTGAATCAGAGTTGGGAGTCTTCCAGCCCA 406

DB 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120

QY 407 GACACCTCTCCGACACCAATGTTGAGGACGACTTCTGATACCCAGGACATCTCG 466

DB 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140

QY 467 GATGAGTTACTGGTAAACATGTTGTTGGCCCACTCCAGATCCCGGACCCCAAGCCCTG 526

DB 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160

QY 527 GCTGTAGCCCTGAGCCCTGCTCAGCCCTGCGGAGCCCGAGCTGGACATCCCACT 586

DB 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180

QY 587 CCCTTCCCAACCTGGGGCCCTCTCAGAACCCACTGAGCGGCTGTTGTTGGCGGGGAA 646

DB 181 PropheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200

QY 647 GAGTGGGAGTTGAGGTGACAGCTTCTACCGGGCGCCCAAGTCTTCCAGCAGACCATC 706

DB 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220

QY 707 TCCTGCCCGAGGGCTCGGCTGGGGTCCGAAGTGGGAGAGCAGGACGCTGCTGGA 766

DB 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240

QY 767 TGGCCAGTCACTCCAGACCTCGGATGCTCCCTGACAGCAGGGGAGTGTAGCTAC 826

DB 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260

QY 827 GTGAGGCATGCTGAGCTGCTGGTGGGGAGTGGCTCTCTGGCGGGCGGGGAGTGG 886

DB 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280

QY 887 CTCTGGGCCCCAGCGCTGGGACATGCCCACATCTGGGAGTGGAGGAGTGTCTC 946

DB 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu 300

QY 947 CCCAACGCGGCGCATGGGCTGATGGCGAGTCCCCAGGACAGGAGGAGGCGGTGTT 1006

DB 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGlyGlyValPhe 320

QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTACGGAAGGAAGCGGACGCTCACA 1066

DB 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340

QY 1067 CGCTATGCCCTCTGGTCTGTGGGGAGTCAATGGGCCCGGAGGAGGAGGAGGAGGAG 1126

DB 341 ArgTyrAlaLeuTrpPheCysValGlySerTrpProGlnAspGlnProTrpThrLys 360

QY 1127 AGGCTCGTGTGTCGCAAGTTGTGCCACGCTGCTCAGGCGCTTGGTGAAGTGGCCCGG 1186

DB 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGlnMetAlaArg 380

QY 1187 GTAGGGGTGCTCTCCTCCCTGGAGAATCTGTGACCTGCACATTTCCACAGCCACCCA 1246

DB 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400

QY 1247 CTCCTCCCTACCTCCGACCAAGTCAAGGCTACCTGAGGAGCTTGGTGGAGGCTATGGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420

QY 1307 TTCAGGGCCCTGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427

RESULT 4
 ABW02028
 ID ABW02028 standard; protein; 427 AA.
 XX AC ABW02028;
 XX AC
 DT 12-FEB-2004 (first entry)
 XX DE Human interferon regulatory factor (IRF3).

XX Human; genetic disease; muscular dystrophy; cystic fibrosis; cytostatic;
 KW scientific research; gene therapy; interferon regulatory factor; IRF3.
 KW
 XX Homo sapiens.

XX Key Location/Qualifiers
 PH Misc-difference 8
 FT /note= "Encoded by NTC"

XX US6639052-B1.
 XX 28-OCT-2003.
 XX 14-OCT-1999; 99US-00417540.
 XX 30-AUG-1995; 95US-0002993P.
 PR 30-AUG-1996; 96US-00705771.
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA;
 XX WPI; 2003-842790/78.
 DR N-PSDB; AAD63572.

XX New isolated protein and nucleic acid molecules, useful for diagnostic
 PT and therapeutic purposes, e.g. for treating genetic diseases such as
 PT muscular dystrophy or cystic fibrosis.

XX Example 5; Fig 5; Opp; English.

XX The invention relates to isolated new isolated protein and nucleic acid
 CC molecules useful for diagnostic and therapeutic purposes. The invention
 CC is for treating genetic diseases such as muscular dystrophy or cystic
 CC fibrosis, and for in vitro purposes related to scientific research,
 CC synthesis of DNA and manufacture of DNA vectors. The invention is useful
 CC in gene therapy. The present sequence is human interferon regulatory
 CC factor (IRF3)

XX Sequence 427 AA;

Alignment Scores:
 Pred. No.: Length: 427
 Score: 2326.00 Matches: 427
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.18% Indels: 0
 DB: 7 Gaps: 0

US-09-975-253-1 (1-1426) x ABW02028 (1-427)

QY 47 ATGGGAACCCCAAGCCAGGNTCCCTGGCTGGTGGTGGAGCTGGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20

QY 107 CTGGAGGGCGTGGCTGGTGAACAAGACCGCACGCGCTTCCGATCCCTTGGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValIleLysSerArgThrArgPheArgIleProTyrLysHis 40
 QY 167 GGCCTACGGCAGGATCACAGCAGGAGGATTTCCGAATCTTCCAGGCTGGGCGGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTGCATATGTTCCCGGAGGATNAGCCAGACTGCCAACCTCGGAAGAGAAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAspPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGTTGCGTTTACGAGGAGCAGGACCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAAATCTACGAGTTTGAACCTCAGGAGTTGGGACTTTTCCAGCCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCGGACACCACTGCTGGAGGAGTACTTCTGATCCCAAGGAGACATTCG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGAGTTACTGGTAAATGTTGTCGCCCTCCACTCCAGATCCGGACCCCAAGCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTCGAGCCCTGCCCTCAGCCCTCGGAGCCCTCAGCTTGGACAATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGCGGCTGTGGTCCCGGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTTCGAGGTGACAGCCCTTCTACCGGGCCGCCCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGGCCCTGCGGCTGGTGGGTCCGAAGTGGGAGACAGACCGCTGCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCCAGACCCCTGGCATGTCCTGACAGACAGGGAGTGTAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGCATGTGCTGAGCTGCTGGGTGGGGACTGGCTCTCTGGCGGGCCGGGAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCCCCAGCGCTGGGCACTGCCACATCTGGGCACTGAGGAGAGTGTCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluLeuLeuLeu 300
 QY 947 CCCAACAGGGGCATGGGCTGATCGCGAGTCCCAAGGACCAAGGAGGAGGCGCTGTTT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACTTTCAGGAGGAGGAGCGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGTTCTGTGTGGGAGTCAATGGCCCCAGGACCGGCTGGACCAAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTCGTATGCTCAAGTTGTGCCACCTGCTCCCTCAGGCGCTTGTGTAAGAAATGCCCGG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGCTCTCTCCCTGGAGAATACTGTGGAGCAATCTGTGGACCTGCACATTTCCAACAGCCACCA 1246

Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCTCGACAGGCTACCTGACGAGCTGTGTGAGGCGCATGGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTyrIysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGGCCCTGGGGAGGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427

RESULT 5
 ADG47651
 ID ADG47651 standard; protein; 427 AA.
 XX
 AC ADG47651;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human IRF3.
 XX
 KW muscular dystrophy; cystic fibrosis; hypertension; angina pectoris;
 KW myocardial infarction; ulcer; asthma; allergy; psychosis; depression;
 KW migraine; vomiting; benign prostatic hypertrophy; osteoporosis; human.
 XX
 OS Homo sapiens.
 XX
 FN US2003208043-Al.
 XX
 PD 06-NOV-2003.
 XX
 PF 04-JUN-2003; 2003US-00453478.
 XX
 PR 30-AUG-1995; 95US-0002993P.
 PR 30-AUG-1996; 96US-00705771.
 PR 14-OCT-1999; 99US-00417540.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Moore PA, Gentz RL, Ji H, Ni J, Hu J;
 XX
 XX WPI; 2003-864796/80.
 DR N-PSDB; ADG47640.
 XX
 XX New human polypeptides and polynucleotides, useful for diagnosing or
 PT treating genetic diseases such as muscular dystrophy or cystic fibrosis,
 PT hypertension, asthma, depression or osteoporosis.
 XX
 PS Claim 1; SEQ ID NO 16; 56bp; English.
 XX

CC The invention relates to an isolated human polypeptide. The polypeptides,
 CC polynucleotides, agonists or antagonist are useful for diagnosing or
 CC treating genetic diseases such as muscular dystrophy or cystic fibrosis,
 CC hypertension, angina pectoris, myocardial infarction, ulcers, asthma,
 CC allergies, psychoses, depression, migraine, vomiting, benign prostatic
 CC hypertrophy or osteoporosis. The polypeptides and polynucleotides are
 CC useful for in vitro purposes related to scientific research, synthesis of
 CC DNA and manufacture of DNA vector. The present sequence represents the
 CC amino acid sequence of human IRF3.
 XX

SQ Sequence 427 AA;

Alignment Scores:
 Pred. No.: 1.9e-166 Length: 427
 Score: 2326.00 Matches: 427
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.18% Indels: 0
 DB: 7 Gaps: 0

US-09-975-253-1 (1-1426) x ADG47651 (1-427)

QY 47 ATGGGAACCCCAAGCCAGGNTCTGCTGGTGTGCGAGCTGGAGCTGGGGCAA 106

Db 1 MetGlyThrProIysProArg***LeuProThrLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGAGGGCGTGGCTGGGTGAACAAGAGCCGACGCGCTTCCGATCCCTTTGAGACAC 166
 Db 21 LeuGluGlyValAlaTyrValAsnLysSerArgThrArgPheArgIleProThrIysHis 40
 QY 167 GSCCTACGGCAGGATGCACAGCAGGAGATTTCCGAATCTTCCAGCCTGGGCCGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnAspPheGlyIlePheGlnAlaTyrAlaGluAla 60
 QY 227 ACTGCTGCATATGTTCCCGGAGGATAGCCAGACCTGCCAACCTGGAAGAGGAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrIysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCCGAAAGAGGTGGTGTAGCAGAGCAGCCGAGCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAATCTACGAGTTTGTGAACCTCAGGAGTGGGACTTTTCCAGCCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACTCTCCGGACACCAATGGTGGAGCGACTTCTGTATACCCAGCAAGACATCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGATTACTGGGTAAACATGGTGTGGCCCTCCAGATCCCGGACCCCAAGCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGGCGAGCCCGCCAGCTTGGACATCCCAT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCTTCCCAAACTGGGGCCCTCTGAGAAACCACTGAAGCGGCTTGTGGTGGCGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTCGAGGTGACAGCTTCTACCGGGCCCGCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGGCTGGCTGGTGGGTCCGAGTGGGAGCAGCAGCCTGCTCGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCCAGACCTGGCATGTCTCCCTGACAGACAGGGGAGTGATGAGCTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGCATGTGCTGAGCTGCTGGTGGGGAGTGGCTCTCTGGCGGGCGGGCAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCCCCAGCGCTGGGCACTGCCACATCTGGCAGTGGCAGTGGAGGAGCTGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu 300
 QY 947 CCCAACAGCGGGCATGGGCTGTGAGTGGCGAGGTCCCGAGGCAAGGAGGAGGCGTGT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTAACCTTACGGAAGAGCGGAGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGGTCTGTGTGGGGAGTCAATGGCCCGCCAGGACCAAGCCGCTGGACCAAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProThrThrIys 360
 QY 1127 AGGCTCGTGTGTCAGAGTTGTGCCACGTCGCTCAGGGCTTGTGTGTAAGAAATGCCCCGG 1186

Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 390

QY 1187 GTAGGGGGTCCCTCCCTGGAGATACCTGGACCTGCACATTTCCACAGCCACCCCA 1246

Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisSerAsnSerHisPro 400

QY 1247 CTCCTCCCTCACCTCCGACCACTACAGGCTTACCTGCAGCACTTGTGGAGGCGATGGAT 1306

Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420

QY 1307 TTCAGGGCCCTGGGAGAC 1327

Db 421 PheGlnGlyProGlyGluSer 427

RESULT 6

AA15102

ID AA15102 standard; protein; 427 AA.

XX AC AA15102;

XX DT 25-JAN-2000 (first entry)

XX DE Modified Interferon Regulatory Factor-3 protein (IRF-3).

XX KW Interferon regulatory factor; IRF-3; transactivation domain; serine; threonine; carboxy terminus; aspartic acid; mutant; phosphorylation; post-translational modification; sendai virus; cancer treatment; herpes; activator; promoter; PRDI/PRDIII; ISRE regulatory element; stimulation; DNA binding; transcriptional activity; transcriptional co-activator; CBP/p300; proteasome mediated degradation; viral infection; influenza; HIV infection; activate; cytokine gene; target cell.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 396 /note= "Wild type Ser replaced with Asp"

FT Misc-difference 398 /note= "Wild type Ser replaced with Asp"

FT Misc-difference 402 /note= "Wild type Ser replaced with Asp"

FT Misc-difference 404 /note= "Wild type Ser replaced with Asp"

FT Misc-difference 405 /note= "Wild type Ser replaced with Asp"

XX WO9951737-Al.

XX PD 14-OCT-1999.

XX PF 07-APR-1999; 99WO-CA000314.

XX PR 07-APR-1998; 98CA-02234588.

XX PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

XX PI Hiscott J, Lin R;

XX DR WPI; 1999-620201/53.

XX DR N-PSDB; AAX90992.

XX PT Carboxy-terminus modified highly active forms of interferon regulatory factor proteins used for the treatment of viral infections.

XX PS Claim 8; Fig 10; 93pp; English.

XX The present protein sequence is the interferon regulatory factor-3 (IRF-3) protein, that is modified in the transactivation domain. The serine and threonine residues in the carboxy terminus are modified post-translationally, by phosphorylation, following sendai virus infection. The modified IRF-3, substituted with aspartic acid, functions as a strong activator of promoters containing PRDI/PRDIII or ISRE regulatory

CC elements, stimulation of DNA binding and transcriptional activity, association of IRF-3 with the transcriptional co-activator CBP/p300 and proteasome mediated degradation. The modified IRF proteins are used for the treatment of viral infections like, influenza, herpes or HIV infection. They may also be used to activate a cytokine gene, in cancer treatment or to modify a target cell of an organism

XX SQ Sequence 427 AA;

Alignment Scores:

Pred. No.: 8.6e-165 Length: 427

Score: 2304.00 Matches: 421

Percent Similarity: 98.59% Conservative: 0

Best Local Similarity: 98.59% Mismatches: 6

Query Match: 85.36% Indels: 0

DB: 2 Gaps: 0

US-09-975-253-1 (1-1426) x AA15102 (1-427)

QY 47 ATGGGAACCCCAAGCCACGNTCCCTGGCTGTGTCGACGTGGACCTGGGCGCA 106

Db 1 MetGlyThrProLysProArgIleLeuProTyrPheValSerGlnLeuAspLeuGlyGln 20

QY 107 CTGAGGGCGTGGCTGGGTGAACAAGAGCGCGCTTCCGCATCCCTTGAAGCAC 166

Db 21 LeuGluGlyValAlaTyrPValAsnLysSerArgThrArgPheArgIleProTyrLysHis 40

QY 167 GGCCTACGGCAGGATGACAGCAGGAGATTTCCGAATCTTCCAGGCTGGGCGGCGCC 226

Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTyrPAlaGluAla 60

QY 227 ACTGCTGCATATGTTCCCGGGAGGATAGGCACACCTGCCCACTGGAGAGAAATTC 286

Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTyrLysArgAsnPhe 80

QY 287 CGCTCTGCCCTCAACCCGCAAGAGGGTTCGTTTAGCAGAGACCGGAGAACGACCT 346

Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100

QY 347 CACGACCCACATAAAATCTACGAGTTTGTGAACCTCAGAGTTGGGAGCTTTTCCAGGCA 406

Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120

QY 407 GACACCTCCGGACACCAATGGTGGAGCAGTACTCTGTATACCCAGAGAACATTCCTG 466

Db 121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspLeu 140

QY 467 GATGAGTTACTGGGTAAACATGTTGTGGCCCACTCCAGATCCGGGACCCCAAGCCCTG 526

Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160

QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGGGAGCCCGGAGCTGGACATCCCACT 586

Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180

QY 587 CCCTTCCCAACCTCGGGCCCTCTCAGAACCCACCTGAAGCGGCTGTGTGGCCGGGAA 646

Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200

QY 647 GAGTGGAGTTCCGAGGTGACAGCCCTTCTACCGGGGCGCCCAAGTCTTCCAGCAGACATC 706

Db 201 GluTyrGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnThrIle 220

QY 707 TCCTGCCGGAGGCGCTCGGCTGGTGGGTCCGAGTGGGAGAGAGGAGCGCTGCTGGA 766

Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240

QY 767 TGGCCAGTCACACTGCCAGACCTGTCATGTCCTGACAGACAGGCGAGTGTAGCTAC 826

Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260

QY 827 GTGAGGCATGTGCTGAGCTGCCCTGGGTGGGAGCTGCGCTCTCTGGCGGCGCGGAGTGG 886

	Db	261	ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp	280
	Qy	887	CTCTGGGCCCCAGCGGTGGGCACACTGCCACACATACTTGGCGAGTGACGAGAGCTGCTC	946
	Db	281	LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrrpAlaValSerGluLeuLeuLeu	300
	Qy	947	CCCACACGGCGGCGATGGCCCTGATGGCGAGGTCCCCAAGGACAAGGAAGAGGCGTGT	1006
	Db	301	ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe	320
	Qy	1007	GACCTGGGGCCCTTCATTGTAGATCTCATTTACCTTCACGCAAGGAACGCGACGCTCACCA	1066
	Db	321	AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro	340
	Qy	1067	CGCTATGCCCTCTGGTTCTGTGTGGGGAGTCAATGGCCCCAGGACCAGCGTGGAACAAG	1126
	Db	341	ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys	360
	Qy	1127	AGGCTCGTGATGGTCAAGGTTGTGCCACCGTGCCTCAGGGCCTTGGTAGAAATGGCCCGG	1186
	Db	361	ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg	380
	Qy	1187	GTAGGGGTGCTCTCTCCCTGGAGAATCTGTGGA CCTGCACTTTCACACGCCACCCA	1246
	Db	381	ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleAspAsnAspHisPro	400
	Qy	1247	CTCTCCCTCACCTCCGACCACTCAAGGCTTACTGCAAGACTTGGTGGAGGCGCATGGAT	1306
	Db	401	LeuAspLeuaspaspaspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp	420
	Qy	1307	TTCAGGGGCCCTGGGGAGAC	1327
	Db	421	PheGlnGlyProGlyGluSer	427

RESULT 7
ADG77016

ID ADG77016 standard; protein; 455 AA.

AC ADG77016;

DT 11-MAR-2004 (first entry)

DE Human nucleic acid associated polypeptide (NAAP) 44.

nucleic acid associated polypeptide; NAAP; cytostatic;
antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
thrombolytic; gene therapy; cell proliferative; cancer; atherosclerosis;
neurological; epilepsy; Huntington's disease; stroke; immune;
inflammatory; AIDS; allergy; developmental; Hypothyroidism;
Cushing's syndrome; infection; human.

OS Homo sapiens.

XX PN WO2003076586-A2.

XX
PD 18-SEP-2003XX
DE
05-MAR-2003.

06-MAR-2002 2002IS-03633309

PR 15-MAR-2002; 2002US-0364438P.
PR 18-APR-2002; 2002US-0373881P.
PR 19-APR-2002; 2002US-0373881P.

XX
DA (TNAV) TNAVET CNAVMT CS TNC

XX

PI Zebarjadian Y, Baughn MR,

PI Lehr-Mason PM, Jin P, Hawkins PR, Wilson AD, Swarnakar A, He A;
PI Hafalia AJA, Tran B, Duggan BM, Emerling BM, Borowsky ML, Yao MG;
PI Chawla NK, Burford N, Khare R, Lee S, Becha SD, Lee SY;

WPI; 2003-756814/71.
N-PSDB; ADG77069.

New human nucleic acid associated proteins (NAAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or infections.

Claim 1: SEO ID NO 44: 373pp: English.

This invention relates to a novel isolated human nucleic acid associated polypeptide (NAAP). The invention may be useful for the development of compounds with a cytostatic, antiarteriosclerotic, anticonvulsant, neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic, anti-inflammatory or thymomimetic activity and also for gene therapy. The invention may prove useful for the development of treatments for diseases or conditions associated with the decreased expression or overexpression of NAAP, such as cell proliferative (for example cancer, atherosclerosis), neurological (for example epilepsy, Huntington's disease, stroke), immune/inflammatory (for example AIDS, allergies) and developmental (for example Hypothyroidism, Cushing's syndrome) disorders, or infections. The present sequence is that of a human NAAP protein of the invention.

Sequence 455 AA:

Alignment Scores:		
Pred. No.:	2,086-164	Length:
Score:	2299.00	Matches:
Percent Similarity:	93.63%	Conservative:
Best Local Similarity:	93.41%	Mismatches:
Query Match:	85.18%	Indels:
DB:	7	Gaps:
		455

US-09-975-253-1 (1-1426) x ADG77016 (1-455)

47	ATGGGAACCCCAAGCCACCGNTCCTCGCTGGCTGGTGTCTGCAGCTGCAGCTGGGGCAA	106
Db		
1	MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGlyGln	20
107	CTGGAGGCGGTGGCTGGGTCAACAAGAGCGCACCGCTTCGCGATCCCTTGGAGCAC	166
Db		
21	LeuGluGlyValAlaIrrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis	40
167	GGCCTACCGCAGGATGCACACGACGACATTTCGGAATCTTCAGACCTTGGGCGCGAGGCC	226
Db		
41	GlyLeuArgGlnAspAlaGlnGlnLysPheGlyIlePheGlnAlaIrrpAlaGluAla	60
227	ACTGTCGCATATGTTCCGGGAGGGATAAGCCAGACCTGCCAACCTGGAGAGGAATTC	286
Db		
61	ThrGlyAlaTyrrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe	80
287	CGCTCTGCCTCAACCGGAAGAAGGTTTCGGTTTACGAGAGACCGGAGCAGAGACCCCT	346
Db		
81	ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro	100
347	CACGACCCACATAAAATCTACGAGTTTGCAATCTCA-----	382
Db		
101	HisaspProHisLysIleTyrrGluPheValAsnSerGlyLeuProAlaTrpGlySerCys	120
382	-----	382
Db		
121	GlyArgLeuGlyProGlnGlyPheLeuAlaLeuLeuLeuProSerProLeuPheSerThr	140
383	GGAGTTGGGACTTTTCCAGCCAGACACTCTCTCCGACACCAATCGTGGAGGCAGTACT	442
Db		
141	GlyValGlyAspPheSerGlnProAspThrSerProAspThrAsnGlyGlyGlySerThr	160
443	TCTGATACCCAGGAGACATTCGTGATGAGTTACTGGTAAACATGTGTGGCCCCCACC	502
Db		
161	SerAspThrGlnGluAspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeu	180
503	CCAGATCCGGACCCCCAAGCTTGGCTGTAGCCCTTCAGCCCTTCGCCCTCAGGCCCTCGGG	562
QY		

Percent Similarity:	98.41%	Conservative:	0
Best Local Similarity:	98.41%	Mismatches:	6
Query Match:	74.77%	Indels:	1
DB:	4	Gaps:	1
US-09-975-253-1 (1-1426) x ABB11963 (1-392)			
QY	212	GCCTGGCGGAGCCACCTGCTGCATATGTTCCCGGAGGGATAAGCCAGACCTGCCAAC	271
Db	15	AlaTTPAlaGluAlaThrGlyAlaTyrValProGlyArgAspLeuProThr	34
QY	272	TGAGAGAGCAATTCCTGCTGCTCCCTCAACCCCAAGAGGGTGTGGTTAGCAGAGGAC	331
Db	35	TrpLysArgAsnPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAsp	54
QY	332	CGGAGCAGGACCTCACCAGCCACATAAATCTACGAGTTTGTGAACTCAGAGTTGGG	391
Db	55	ArgSerLysAspProHisAspProHisLysLysLysLysLysLysLysLysLysLysLys	74
QY	392	GACTTTTCCAGCCAGACACCTCTCCGGACACCAATGGTGGAGGAGTACTTCTGATACC	451
Db	75	AspPheSerGlnProAspThrSerProAspThrAsnGlyGlySerThrSerAspThr	94
QY	452	CAGAGAGACATCTCGATGAGTTACTGGGTAACATGCTGTGGCCCACTCCCGAGATCCG	511
Db	95	GlnGluAspLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspPro	114
QY	512	GGACCCCAAGCCTGCTGTAGCCCTGAGCCCTGACCCCTGCGGAGCCGAGCCAGC	571
Db	115	GlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuArgSerProSer	134
QY	572	TTGGCAATCCACTCCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAGCGGCTG	631
Db	135	LeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu	154
QY	632	TTGGTCCGGGAGAGTGGAGTTCGAGTGCACAGCCTTCTACGGGGCCGCCAAGTC	691
Db	155	LeuValProGlyGluGluTrpGluPheGluValThrAlaPheLysArgGlyVal	174
QY	692	TTCCAGCAGACCACTCTCTGCGGAGGCGCTGCGGTGGTGGTCCGAGTGGGAGAC	751
Db	175	PheGlnGlnThrLysSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp	194
QY	752	AGACGCTGCTGATGCGGATGCTGACACTGCCAGACCTGCGATGCTCCCTGACAGCAGG	811
Db	195	ArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArg	214
QY	812	GGAGTGATGAGTACGTGAGGAGTGTGCTGAGCTGCTGCTGGTGGGAGTCTCTCTGG	871
Db	215	GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrp	234
QY	872	CGGCGCGGAGTGGCTCTGGGCGGAGCGGCTGCGGAGTGGGAGTGGGAGTGGGAGT	931
Db	235	ArgAlaGlyGlnTrpLeuTrpAlaGluArgLeuGlyHisCysHisThrTyrTrpAlaVal	254
QY	932	AGCAGAGAGTCTCTCCCAACAGCGGCGATGGGCTGATGCGGAGTCCCGCAGGACAG	991
Db	255	SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys	274
QY	992	GAAGGAGCGTGTGGTACCTGGGCGCTTCTTCTGTA-----GATCTG	1033
Db	275	GluGlyGlyValPheAspLeuGlyProPheLysValGlySerLeuGlyProProAspLeu	294
QY	1034	ATTACCTTCAGGAGGAGGAGCGCTCACAGCTATGCTGCTGCTGCTGCTGCTGCTG	1093
Db	295	IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGly	314
QY	1094	GAGTCATGCGCCAGGACCCGCTGACAGAGGCTCGATGCTGCTGCTGCTGCTGCTG	1153
Db	315	GluSerTrpProGlnAspGlnProTrpThrLysArgLeuValMetValLysValValPro	334
QY	1154	ACGTGCTTCAGGGCTCTGGTAGAAATGGCCCGGCTAGGGGTGCTCTCTCTGCTGAGAT	1213

Db	335	ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn	354
QY	1214	ACTGTGGACCTGCACATTTCCAAACAGCCACCTCTCCCTCACCTCCGACGACGACG	1273
Db	355	ThrValAspLeuHisLeuSerAsnSerHisProLeuSerLeuThrSerAspGlnTyrLys	374
QY	1274	GCCTACCTGCAGGACTTGTGTGGAGGCGCATGTTTCCAGGGCCCTGGGGAGAGC	1327
Db	375	AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer	392
RESULT 9			
AAM79970			
ID	AAM79970 standard; protein; 392 AA.		
XX	AAM79970;		
AC	AAM79970;		
XX	06-NOV-2001 (first entry)		
DT	Human protein SEQ ID NO 3616.		
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation.		
XX	Homo sapiens.		
OS	WO200157190-A2.		
FN	09-AUG-2001.		
XX	05-FEB-2001; 2001WO-US004098.		
PF	03-FEB-2000; 2000US-00496914.		
XX	27-APR-2000; 2000US-00560875.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
XX	(HYSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;		
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AU, Yang Y, Wejhrman T, Goodrich R;		
XX	WPI; 2001-476283/51.		
DR	N-PSDB; AAKS3103.		
XX	Nucleic acids encoding polypeptides with cytokine-like activities, useful		
PT	in diagnosis and gene therapy.		
XX	Claim 20; Page 397; 6221pp; English.		
PS	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
XX	encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111		
CC	(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the		
CC	sequence listing were missing at the time of publication		
XX	Sequence 392 AA;		
SQ	Alignment Scores:		

Pred. No.:	2,91e-143	Length:	392
Score:	2018.00	Matches:	372
Percent Similarity:	98.41%	Conservative:	0
Best Local Similarity:	98.41%	Mismatches:	0
Query Match:	74.77%	Indels:	6
DB:	4	Gaps:	1
US-09-975-253-1 (1-1426) x AM79970 (1-392)			
QY	212	GCCTGGCGGAGCCACTGGTGCATATGTTCCGGGAGGAGTAAGCCAGACCTGCCAAC	271
Db	15	AlaTrpAlaGluAlaThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThr	34
QY	272	TGGAGAGCAATTCGGCTCGCTCCCTCAACCCCAAGAGGGTGGTTAGCAGAGAC	331
Db	35	TrpLysArgAspPheArgSerAlaLeuAsnArgLysGluGlyLeuAlaGluAsp	54
QY	332	CGGAGCAGGACCCCTCAGCACCACATAAAATCAGAGTTTGTGAATCAGGAGTTGG	391
Db	55	ArgSerLysAspProHisAspProHisLysLysLysLysLysLysLysLysLysLys	74
QY	392	GACTTTTCCAGCAGACCTCTCCGACACCAATGTGGAGGAGTACTTCTGTATCC	451
Db	75	AspPheSerGlnProAspThrSerProAspThrAsnGlyGlySerThrSerAspThr	94
QY	452	CAGGAAGACATCTGGATGAGTACTGGGTAACTAGTGTGGTGGCCCACTCCAGATCG	511
Db	95	GlnGluAspLysLeuAspLysLeuGlyAsnMetValLeuAlaProLeuProAspPro	114
QY	512	GGACCCCAAGCCTGGCTGTAGCCCTGAGCCCTGACCCCTGAGCCCTGGAGCCCGC	571
Db	115	GlyProSerLeuAlaValAlaProGluProCysProGlnProLeuArgSerProSer	134
QY	572	TTGGACATCCACCTCCCTCCCAACCTGGGGCCCTCTGAGAACCACTGAAGCGGTG	631
Db	135	LeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu	154
QY	632	TTGGTCCGGGGAGAGTGGAGTTCGAGGTGACAGCTTCTACGGGGCCGCAAGTC	691
Db	155	LeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnVal	174
QY	692	TTCCAGCAGACATCTCTCCCGGAGGCGCTGGCTGGTGGTCCGAACTGGAGAC	751
Db	175	PheGlnGlnThrLysSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp	194
QY	752	AGGAGCTGCTGGATGCGCAGTCACTGCGCAGACCTGGCATGTCCTGACACAGG	811
Db	195	ArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArg	214
QY	812	GGAGTGATGAGTACGTGAGGCATGCTGAGCTGCCTGGTGGGGGACTGCTCTCTGG	871
Db	215	GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrp	234
QY	872	CGGCGCGGCGAGTGGCTCTGGCGCCAGCGGCTGGGGCTGCGACACATCTGGGCACTG	931
Db	235	ArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaVal	254
QY	932	AGCGAGGAGTCTCTCCCAACAGCGGGGATGGGCTGTATGGCGAGGTCCCAAGGACAAG	991
Db	255	SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys	274
QY	992	GAAGGAGCGTGTGTGACCTGGGGCCCTCATTTGTA-----GATCTG	1033
Db	275	GluGlyGlyValPheAspLeuGlyProPheIleValGlySerLeuGlyProProAspLeu	294
QY	1034	ATTACCTTCACGAGGAGGAGCGCTCACACGCTATGCGCTCTGCTGTGTGTGGG	1093
Db	295	IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGly	314
QY	1094	GAGTCATGCGCCAGGACCGCGTGGACCAAGAGGCTGTGATGGTCAAGGTTGTGCC	1153
Db	315	GluSerTrpProGlnAspGlnProTrpThrLysArgLeuValMetValLysValPro	334

QY	1154	ACGTGCTCAGGCGCTTGTGTAGAAATGGCCCGGTAGGGGTGCTCTCTCCCTGGAGAT	1213
Db	335	ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn	354
QY	1214	ACTGTGGACCTGCACATTTCCACACCCACCTCCCTCCTCCTCCTCCACGAGTACAAG	1273
Db	355	ThrValAspLeuHisLysSerAsnSerHisProLeuSerLeuThrSerAspGlnTyrLys	374
QY	1274	GCCTACTCAGGACTTGTGGAGGCGATGGATTTCCAGGGCCCTGGGAGAGC	1327
Db	375	AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer	392
RESULT 10			
AAV15104	AAV15104 standard; protein; 542 AA.		
ID	XX		
AC	AAV15104;		
DT	25-JAN-2000 (first entry)		
DE	Chimeric protein IRF-7(1-246)/IRF-3(132-427).		
KW	Interferon regulatory factor; IRF; chimeric protein; serine; threonine;		
KW	carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;		
KW	PRDI/PRDIII; promoter; ISRE regulatory element; stimulation; herpes;		
KW	activator; DNA binding; transcriptional activity; viral infection;		
KW	proteasome mediated degradation; influenza; HIV infection; cytokine gene;		
KW	target cell.		
OS	Homo sapiens.		
XX	Synthetic.		
Key	Location/Qualifiers		
Region	1..246		
FT	/note= "Corresponds to modified amino-terminal domain of		
FT	IRF-7 protein"		
Region	247..542		
FT	/note= "Corresponds to modified carboxy-terminus of IRF-3		
FT	protein"		
Misc-difference	511		
FT	/note= "Wild type Ser replaced with Asp"		
Misc-difference	513		
FT	/note= "Wild type Ser replaced with Asp"		
Misc-difference	517		
FT	/note= "Wild type Ser replaced with Asp"		
Misc-difference	519		
FT	/note= "Wild type Ser replaced with Asp"		
Misc-difference	520		
FT	/note= "Wild type Ser replaced with Asp"		
XX	WO9951737-A1.		
XX	14-OCT-1999.		
XX	07-APR-1999; 99WO-CA000314.		
XX	07-APR-1998; 98CA-02234588.		
XX	(DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.		
XX	Hiscott J, Lin R;		
XX	WPI; 1999-620201/53.		
XX	N-PSDB; AAX90994.		
XX	Carboxy-terminus modified highly active forms of interferon regulatory		
XX	factor proteins used for the treatment of viral infections.		
XX	Claim 11; Fig 13; 93pp; English.		
XX	The present sequence is the chimeric protein comprising, residues 1-246		
CC	from the modified amino-terminal domain of interferon regulatory factor,		

CC IRP-7 and 132-427 residues from the carboxy-terminus of modified IRP-3
 CC (5D) protein. The serine and threonine residues are post-translationally
 CC modified by phosphorylation, following sendai virus infection. The
 CC modified IRP, substituted with aspartic acid, functions as a strong
 CC activator of promoters containing ISRE and PRDII regulatory
 CC elements, stimulation of DNA binding and transcriptional activity and
 CC proteasome mediated degradation. The modified IRP sequences are used for
 CC the treatment of viral infections like, influenza, herpes or HIV
 CC infection. They may also be used to activate a cytokine gene, in cancer
 CC treatment or to modify a target cell of an organism
 XX
 SQ Sequence 542 AA;

Alignment Scores:
 Pred. No.: 2,79e-124 Length: 542
 Score: 1766.50 Matches: 344
 Percent Similarity: 67.73% Conservative: 17
 Best Local Similarity: 64.54% Mismatches: 61
 Query Match: 65.45% Indels: 111
 DB: 2 Gaps: 6

US-09-975-253-1 (1-1426) x AAY15104 (1-542)

QY	62	CCACGGNTCCG-----CCCTGGCTGGTGTGCGACCTGGGCACTGGAGGCG	115
DB	10	ProArgValLeuPheGlyGluTrpLeuLeuGlyGluLeuSerSerGlyCysTyrGluGly	29
QY	116	GTGGCTGGTGAACAGAGCGCAGCGCTCCGCATCCCTTGAAGCAGCGCTACGG	175
DB	30	LeuGlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpIshisPheAlaArg	49
QY	176	CAGGATGCACAGCAGGAGATTTCCGAATCTTCCAGGCTGGCCGAGGCACTGTGCA	235
DB	50	LysAspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArg	69
QY	236	TATGTTCCC-----GGAGGATAACCCAGACCTGCCAACCC-----	271
DB	70	TrpProSerSerArgGlyGlyGlyProProProGluAlaGluTrpAlaGluArgAla	89
QY	272	---TGAAGAGGAATTTCCGCTCTGCGCTCAACCGCAAGAGGTTGCGTTTACGAGAG	328
DB	90	GlyTrpLysThrIshisPheArgCysAlaLeuArgSerThrArgArgPheValMetLeuArg	109
QY	329	GACCGGACGAGGACCTCCAGCCACATATAATCTACGAGTTGTG-----	376
DB	110	AspAsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCys	129
QY	377	---AACTCAGGAGTTGG-----GACTTTTCCCGCCAGACACCTCTCCGGAC	421
DB	130	TrpArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaAlaValProPro	149
QY	422	ACCAATGGTGA-----	433
DB	150	ProGlnGlyGlyProProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaPro	169
QY	433	-----	433
DB	170	GlyProLeuProAlaProAlaGlyAspLysGlyAspLeuLeuLeuGlnAlaValGlnGln	189
QY	433	-----	433
DB	190	SerCysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLys	209
QY	433	-----	433
DB	210	AlaProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGly	229
QY	434	-----GGCAGTACTCTCGAT	448
DB	230	LeuProAlaGlyGluLeuTyrGlyTrpAlaValGluThrThrProSerProThrSerAsp	249
QY	449	ACCAGGAGACATTTCTGATCAGTACTTGGTAACTGGTGTGCGCCCACTCCACAT	508

Db	250	ThrGlnGluAspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuLeuProAsp	269
QY	509	CCGGAGCCCCCAACCTGGCTGAGCCCTCAGCCCTGCGCTCAGCCCTCGGAGCCCC	568
Db	270	ProGlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuArgSerPro	289
QY	569	AGCTTGGACAATCCACTCCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGG	628
Db	290	SerLeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArg	309
QY	629	CTGTTGGTCCGGGGAGAGAGTGGAGTTCAGGTGACAGGCTTCTACCGGGCCGCCAA	688
Db	310	LeuLeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGln	329
QY	689	GTCTTCCAGCAGACCATCTCTGCCCGAGCGCTGCGGCTGGTGGGGTCCGAAGTGGGA	748
Db	330	ValPheGlnGlnThrIshisPheCysProGluGlyLeuArgLeuValGlySerGluValGly	349
QY	749	GACAGCAGCTGCTGGTGGATGGCCACTCACCTGCCAGACCCCTGGCATGTCCCTGACAG	808
Db	350	AspArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAsp	369
QY	809	AGGAGAGTATGAGTACGTAGGCGATGTGCTGAGCTGCTGGTGGGGGAGCTGCTCTC	868
Db	370	ArgGlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeu	389
QY	869	TGCGGGCGGGCAGTGGCTCTGGGCCCGAGCGCTGGGGCAGCTGCCACATACCTGGCA	928
Db	390	TrpArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAla	409
QY	929	GTGAGCAGGAGCTGCTCCCCAACAGCGGCGATGGGCTGTGAGTGGGAGGTCCCAAGGAC	988
Db	410	ValSerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAsp	429
QY	989	AGGAAAGAGGCGTGTTCACCTGGCGCCCTTCATTGTAGTCTGATTACCTTCACGGA	1048
Db	430	LysGluGlyGlyValPheAspLeuGlyProPheIleValAspLeuIleThrPheThrGlu	449
QY	1049	GGAAGCGGACGCTCACACGCTATGCGCTCTGTTCTGTGGGGGAGTCAATGCGCCAG	1108
Db	450	GlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGln	469
QY	1109	GACACGCGCTGGACCAAGAGGCTCGTGTGATGGTCAAGTGTGTGCCCGCTCGCTCAGG	1168
Db	470	AspGlnProTrpThrLysArgLeuValMetValLysValValProThrCysLeuArgAla	489
QY	1169	TTGGTAAATGCGCGGTAGGGGTGCTCTCCCTCGAGAAATCTGTGGACCTGCAC	1228
Db	490	LeuValGluMetAlaArgValGlyAlaSerSerLeuGluAsnThrValAspLeuHis	509
QY	1229	ATTTCACACAGCCACCTCTCCCTCACCTCCGACAGTACAGGCTTACCTGCGAGGAC	1288
Db	510	IleAspAsnAspHisProLeuAspLeuAspAspGlnTyrLysAlaTyrLeuGlnAsp	529
QY	1289	TTGGTGGAGCGCATGGATTTCCAGGGCCCTGGGGAGAGC	1327
Db	530	LeuValGluGlyMetAspPheGlnGlyProGlyGluSer	542

RESULT 11
 AAG00953
 ID AAG00953 standard; protein; 102 AA.
 AC AAG00953;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5034.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX
 OS gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX

PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 DR N-PSDB; AAC00959.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 PS
 PS Claim 13; SEQ ID NO 5034; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX
 SQ Sequence 102 AA;
 Alignment Scores:
 Pred. No.: 1,468-32 Length: 102
 Score: 546.00 Matches: 100
 Percent Similarity: 98.04% Conservative: 0
 Best Local Similarity: 98.04% Mismatches: 2
 Query Match: 20.23% Indels: 0
 DB: 3 Gaps: 0
 US-09-975-253-1 (1-1426) x AAG00953 (1-102)
 QY 47 ATGGGAACCCCAAGCCACGNTCTCGCTGGTGTGCGAGTGGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGGAGGCGTGGCTGGTGAACAGAGCCGACGCGCTCCCGATCCCTTGGAGAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACCGCAGGATGCACAGCAGGAGATTTCGAATCTCCAGCGCTGGCCGAGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTGCATATGTTCCCGGGAGGGATAGCGACACCTGCCAAGAGGGAATTC 286
 Db 61 ThrGlyAlaTyrrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGGTTGGTTTACGAGGACCGGAGCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLys***Pro 100
 QY 347 CACGAC 352
 Db 101 HisAsp 102
 RESULT 12
 ADD01173

ADD01173 standard; protein; 488 AA.
 ADD01173;
 01-JAN-2004 (first entry)
 Human nucleic acid-associated protein NAAP-11 SEQ ID NO:11.
 human; nucleic acid-associated protein; NAAP; cytostatic;
 antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
 antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
 antiinflammatory; ophthalmological; thymimetic; antiarthritic;
 hepatotropic; antibacterial; virucide; protozoicide; antiparasitic;
 fungicide; gene therapy; cell proliferative disease; cancer;
 atherosclerosis; hepatitis; neurological disorder; Parkinson's disease;
 Alzheimer's disease; stroke; epilepsy; developmental disorder;
 renal tubular acidosis; anaemia; glaucoma; hypothyroidism;
 autoimmune disorder; inflammatory disorder; AIDS; allergy;
 atopic dermatitis; arthritis; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO2003054219-A2.
 XX
 PD 03-JUL-2003.
 XX
 PF 18-DEC-2002; 2002WO-US041115.
 XX
 PR 19-DEC-2001; 2001US-0343004P.
 PR 11-JAN-2002; 2002US-0347633P.
 PR 25-JAN-2002; 2002US-0351749P.
 PR 22-FEB-2002; 2002US-0359498P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Azinzai Y, Baughn MR, Becha SD, Borowsky ML, Burford N;
 PI Elliott VS, Emerling BW, Forsythe IJ, Gorvad AE, Griffin JA;
 PI Kabie AE, Khare R, Lal PG, Lee EA, Lee SY, Li JX, Marquis JP;
 PI Ramkumar J, Richardson TW, Sprague WW, Swarnakar A, Tang YT;
 PI Chawla NK, Warren BA, Yue H;
 DR WPI; 2003-559157/52.
 DR N-PSDB; ADD01230.
 XX
 PT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
 PT infections.
 PT
 PS Claim 1; SEQ ID NO 11; 405pp; English.
 XX
 CC The present invention describes human nucleic acid-associated proteins
 CC designated NAAP-1 to NAAP-57. The human NAAPs have cytostatic,
 CC antiarteriosclerotic, anti-HIV, antiallergic, cerebroprotective,
 CC antiparkinsonian, anticonvulsant, nootropic, neuroprotective,
 CC antiinflammatory, ophthalmological, thymimetic, antiarthritic,
 CC hepatotropic, antibacterial, virucide, protozoicide, antiparasitic and
 CC fungicide activities, and can be used in gene therapy. The NAAP protein
 CC and polynucleotide sequences can be used in diagnosing, treating and
 CC preventing diseases or conditions associated with the decreased
 CC expression or overexpression of NAAP, such as cell proliferative diseases
 CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
 CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS, allergies,
 CC atopic dermatitis, arthritis) and infections (e.g. bacterial, viral,
 CC parasitic, protozoal, fungal). The present sequence represents human NAAP
 CC -11, from the present invention.
 XX
 SQ Sequence 488 AA;
 Alignment Scores:
 Pred. No.: 3,03e-29 Length: 488

Db	306	GlnLeuLeuAspValLeuAspArgGlyLeuIleLeuGlnLeuGlnGlyIleAspLeuYr	325
Qy	893	GGCCAGCGCTGGGGCACTGGCCACACATACTGG-----GCATGAGCGCAGGAG	940
Db	326	AlaIleArgLeuGlyGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp	345
Qy	941	CTGCTCCCAACAGCGGGCATGGGCTCATGGCGAGGTCCCACAGGACAAGGAGGAGGC	1000
Db	346	SerCysProAsn-----ProlleGlnArgGluValLysThrLys	358
Qy	1001	GTGTTTGACCTGGGCGCCCTTCATTGTAGATCTGATTACCTTCACGGAAGGA--AGCGGA	1057
Db	359	LeupheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn	378
Qy	1058	CGCTCACCAACGCTATGCCCTCTGGTTCTGTGGGGAGTCTATGGCCCCAGGACCAAGCCG	1117
Db	379	ThrProProPheGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro	398
Qy	1118	TGGACCAAGAGGCTGTGATGGTCAAGTTGTGCCACGTGCTCAGGCGCTGGTAGAA	1177
Db	399	ArgGluLysLysLeuIleThrValGlnValValProValAlaGlnAlaArgLeuLeuGlu	418
Qy	1178	ATGCCCCGGTAGGGGGTCCCTCCTCCCTGGAGATACCTGTGGACCTGCACATTTCCAAC	1237
Db	419	MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn	437
Qy	1238	AGGCA-----CCCACTCTCCCTCAC-----	1257
Db	438	--ProAspLeuLysAspArgMetValGluGlnPheLysGluLeuHisHisIleTrpGlnS	457
Qy	1258	-----CTCCGACAGTACAGGCCCTACTGCAGGACTTGTGGAGGGCATGGAT	1306
Db	457	erGlnGlnArgLeuGlnProValAlaGlnAlaProGlyAlaGlyLeuGlyValGly	476
Qy	1307	TTCACGGCCCTCG 1320	
Db	477	--GlnGlyProTrp 480	
RESULT 13			
ADJ75465			
ID	ADJ75465	standard; protein; 516 AA.	
XX	ADJ75465;		
AC			
XX			
DT	20-MAY-2004	(first entry)	
XX			
DE	Marker gene related amino acid sequence SEQ ID NO:717.		
XX	bronchial asthma; chronic obstructive pulmonary disease;		
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;		
KW	gene therapy; marker.		
XX			
OS	Homo sapiens.		
XX			
FN	EPI394274-A2.		
XX			
PD	03-MAR-2004.		
XX			
PF	04-AUG-2003; 2003EP-00254857.		
XX			
PR	06-AUG-2002; 2002JP-00229312.		
FR	20-MAR-2003; 2003JP-00077212.		
XX			
PA	(GENO-) GENOX RES INC.		
XX			
PI	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;		
XX			
DR	WPI; 2004-193155/19.		
XX			
PT	Testing for bronchial asthma or chronic obstructive pulmonary disease by		
PT	comparing the expression level of a marker gene in a biological sample		
PT	from a subject with the expression level of the gene in a sample from a		

Testing for bronchial asthma or chronic obstructive pulmonary disease by comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a

healthy subject.

Example 11; SEQ ID NO 717; 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiallergic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 516 AA;

Alignment Scores:
 Pred. No.: 1-27e-25 Length: 516
 Score: 456.50 Matches: 145
 Percent Similarity: 40.15% Conservative: 65
 Best Local Similarity: 27.72% Mismatches: 202
 Query Match: 16.91% Indels: 111
 DB: 8 Gaps: 20

US-09-975-253-1 (1-1426) x ADU75465 (1-516)

PT healthy subject.
 XX Example 11; SEQ ID NO 717; 241pp; English.
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (I) has respiratory
 CC and antiallergic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 516 AA;
 Alignment Scores:
 Pred. No.: 1-27e-25 Length: 516
 Score: 456.50 Matches: 145
 Percent Similarity: 40.15% Conservative: 65
 Best Local Similarity: 27.72% Mismatches: 202
 Query Match: 16.91% Indels: 111
 DB: 8 Gaps: 20
 US-09-975-253-1 (1-1426) x ADU75465 (1-516)
 QY 5 CCAGCTGCC-----CGCAGCGCCCGACCTTCATCGTAGCGCGGACCATGGGAACCCCA 58
 DB 7 ProAlaAlaGlyProAspSerProArgPro-----GlyThrArgAlaAlaPro 23
 QY 59 AGCCACCGNTCTCGCCCTGGTGTGTGCGAGCTGGACCTGGGCAACTGGAGGGCGRG 118
 DB 24 ArgValLeuPheGlyGluTrpLeuLeuGlyGluIleSerSerGlyCysTrpGluGlyLeu 43
 QY 119 GCCTGGGTGAACAAGACGCGCGCTTCCGCATCCCTTGAAGACAGCGCTCGAGCAG 178
 DB 44 GlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpPheHisPheAlaArgLys 63
 QY 179 GATGACACAGCAGAGGATTCGGAATTCCTCCAGGCTGGCGCGGACGCTGGTGCATAT 238
 DB 64 AspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArgTrp 83
 QY 239 GTTCCC-----GGGAGGATAGCCAGACCTGCCAAC----- 271
 DB 84 ProProSerSerArgGlyGlyGlyProProGluAlaGluThrAlaGluAlaGly 103
 QY 272 TGAAGAGAAATTCGCTCTCCCTCAACCCCAAGAGAGGTTCGGTTTACAGAGGAC 331
 DB 104 TrpLysThrAsnPhaArgCysAlaLeuArgSerThrArgPheValMetLeuArgAsp 123
 QY 332 CGAGCAAGGACCCCTCAGACCCACATAAATCTACGAGTTGTG----- 376
 DB 124 AsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCysTrp 143

QY 377 AACTCAGGAGTTGG-----GACTTTTCCAGCCAGACACCTCTCCGACACC 424
 DB 144 ArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaAlaValProPro 163
 QY 425 AATGCTGGA----- 433
 DB 164 GlnGlyGlyProGlyProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaProGly 183
 QY 434 -----GSCAGTACTCTGAT-----ACCCAGGAA 457
 DB 184 ProLeuProAlaProAlaGlyAspGluGlyAspLeuLeuGlnAlaValGlnGlnSer 203
 QY 458 GACATCTCGATGAGTTACTG---GGTAACATGGTGTGGCCCTCCCTCCCA 505
 DB 204 CysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLysAla 223
 QY 506 -----GATCCGGGACCC 517
 DB 224 ProGlyGluGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGlyLeu 243
 QY 518 CCA-----AGCCTGGCTGTAGCCCTGAGCCCTGAGCCCTCCCTAGCCC 556
 DB 244 ProAlaGlyGluLeuTyrglyTrpAlaValGluThrProSerProGlyProGlnPro 263
 QY 557 CTGGGAGCCCGAGCTTGACAAATCCCACTCCC-----TTCCCAAC 598
 DB 264 AlaAlaLeuThrThrGlyGluAlaAlaAlaProGluSerProHisGlnAlaGluProTyr 283
 QY 599 CTGGGCGCTCTGAGAACCCACTGAAGCGGTGTGGTCCGGGGGAAAGTGGAGTTC 658
 DB 284 LeuSerProSerProSerAlaCysThrAlaValGlnGluPro---SerProGlyAlaLeu 302
 QY 659 GAGGTGACAGCCTTCTACCGGGCGCGCAAGTCTTCCAGCAGACCATCTCTCCCGGAG 718
 DB 303 AspValThrIleMetTyrlsGlyArgThrValLeuGlnLysValValGlyHisProSer 322
 QY 719 GGCCTCGCGCTGGTGGG---TCCGAAGTGGGAGACAGCAGCTGCTGGATGGCCAGTC 775
 DB 323 CysThrPheLeuTyrglyProProAspProAlaValArgAlaThrAspProGlnVal 342
 QY 776 ACAGTCCAGACCCCTGGCATGTCCCTGACAGACAGGGAGTGTAGTACGTAGGAGCAT 835
 DB 343 AlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrThrGluGlu 361
 QY 836 GTGCTGAGCTGCTGGTGGGGAGCTGCTCTCGCGGGCGGCGAGCTGCTGGATGGCCAGTC 895
 DB 362 LeuLeuArgHisValAlaProGlyLeuHisGluLeuArgGlyProGlnLeuTrpAla 381
 QY 896 CAGCGCTGGGCACTGCCACACATACCTGGGAGTGGAGGAGCTGCTCCCAACAGC 955
 DB 382 ArgArgMetGlyLysCysLysValTyrTrpGluValGly----- 394
 QY 956 GGGCATGGCGCTGATGGGAG-----GTCCCCAAGAGAC 988
 DB 395 -----GlyProProGlySerAlaSerProSerThrProAlaCysLeuLeuProArgAsn 412
 QY 989 AAGGAAGGAGGCGTGTGTTGACCTGGGCGCTTCTGTTAGATCTGATTACCTTCCAGGAA 1048
 DB 413 CysAspThrProIlePheAspPheArgValPhePheArgGluLeuValGluPheArgAla 432
 QY 1049 GGAAGCGGACGC---TCACCACGCTATCCCTCTGTTGTGGGGAGTGCATGGGCC 1105
 DB 433 ArgGlnArgArgGlySerProArgTyrThrIleTyrLeuGlyPheGlyGlnAspLeuSer 452
 QY 1106 CAGGACAGCGGTGGACCAAGAGGCTCTGATGTGTCAAGGTGTGGCCACGCTGCTCAGG 1165
 DB 453 AlaGlyArgProLysGluLysSerLeuValLeuValLysLeuGluProTrpLeuCysArg 472
 QY 1166 GCCTTGGTAGAAATGGCCCGGTAGGGGTCCCTCCCTCCCTGGAG---AATACTGTGAC 1222
 DB 473 ValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerSerLeuSer 492

QY 1223 CTGCACATTTCCACAGCCACCCACTCTCCCTACCTCGACACAGTACAGGCCTACCTG 1282
 DB 493 LeuCysLeuSerSerAlaAsnSerLeu-----TyrAspAspIleGluCysPheLeu 509

QY 1283 CAGGACTTG 1291
 DB 510 MetGluLeu 512

RESULT 14
 ADJ75413
 ID ADJ75413 standard; protein; 516 AA.

XX AC ADJ75413;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:665.

XX bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker.

XX Homo sapiens.

XX EP1394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003BP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.

XX Example 11; SEQ ID NO 665; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (i) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for

CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX SQ Sequence 516 AA;

Alignment Scores:
 Pred. No.: 1.27e-25 Length: 516
 Score: 456.50 Matches: 145
 Percent Similarity: 40.15% Conservative: 65
 Best Local Similarity: 27.72% Mismatches: 202
 Query Match: 16.91% Indels: 111
 DB: 8 Gaps: 20

US-09-975-253-1 (1-1426) x ADJ75413 (1-516)

QY 5 CCAGCTGCC-----CGCACGCCCGACCTTCCATCGTAGCCGACCATGGAACCCCA 58
 DB 7 ProAlaAlaGlyProAspSerProArgPro-----GlyThrArgAlaAlaPro 23
 QY 59 AAGCCACCGNTCCTCCCTGGCTGGTGTGCGACCTGGACCTGGGGCACTGGAGGCGTG 118
 DB 24 ArgValLeuPheGlyGluTrpLeuLeuGlyGluLeuSerSerGlyCysTyrGluGlyLeu 43
 QY 119 GCCTGGGTGAACAAGAGCGCGCGCTTCCGATCCCTTGAAGCAGCGCCCTACGGCAG 178
 DB 44 GlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpLysHisPheAlaArgLys 63
 QY 179 GATGCACAGCAGGAGATTTCGGAATCTCCAGGCTGGCGGACGAGCCCTGCTGATAT 238
 DB 64 AspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArgTrp 83
 QY 239 GTTCCC-----GGGAGGGATAAGCCAGACCTGCCAAC----- 271
 DB 84 ProProSerSerArgGlyGlyGlyProProGluAlaGluThrAlaGluArgAlaGly 103
 QY 272 TGAAGAGGAATTTCCGCTCTGCCCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGAC 331
 DB 104 TrpLysThrAsnPheArgCysAlaLeuArgSerThrArgArgPheValMetLeuArgasp 123
 QY 332 CGAGCAGAGGACCTCACGACCCACATATAATCTACGAGTTTGTG----- 376
 DB 124 AsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCysTrp 143
 QY 377 AACTCAGAGGTGGG-----GACTTTTCCAGCGACACACCTCTCGGACACC 424
 DB 144 ArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaValProPro 163
 QY 425 AATGGTGA----- 433
 DB 164 GlnGlyGlyProProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaProGly 183
 QY 434 -----GGCAGTACTTCTGAT-----ACCCAGGAA 457
 DB 184 ProLeuProAlaProAlaGlyAspGluGlyAspLeuLeuLeuGlnAlaValGlnSer 203
 QY 458 GACATTCTGGATGAGTTACTG---GGTAACATGGTGTGGCCCACTCCCA----- 505
 DB 204 CysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLysAla 223
 QY 506 -----GATCCGGGAGCCC 517
 DB 224 ProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGlyLeu 243
 QY 518 CCA-----AGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGCTCAGGCC 556
 DB 244 ProAlaGlyGluLeuTyrGlyTrpAlaValGluThrThrProSerProGlyProGlnPro 263
 QY 557 CTGGGAGCCCCAGCTTGGACAATCCCACTCC-----TTCCCAAC 598
 DB 264 AlaAlaLeuThrThrGlyGluAlaAlaAlaProGluSerProHisGlnAlaGluProTrp 283
 QY 599 CTGGGGCCCTCTGAGACCCACTGAAGCGGCTGTGTGGTGGCGGGAAGAGTGGAGTTC 658

QY	377	AACTCAGAGTGTGG-----GACTTTTCCAGCCAGACACCTCTCGGACACC	424
DB	144	ArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaValProPro	163
QY	425	AATGGTGA-----	433
DB	164	GlnGlyGlyProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaProGly	183
QY	434	-----GGCAGTACTTCTGTAT-----ACCCAGGAA	457
DB	184	ProLeuProAlaProAlaGlyAspGluGlyAspLeuLeuLeuGlnAlaValGlnSer	203
QY	458	GACATCTCGATGAGTTACTG---GGTAACATGGTGTGGCCCACTCCCA	505
DB	204	CysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLysAla	223
QY	506	-----GATCCGGGACCC	517
DB	224	ProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGlyLeu	243
QY	518	CCA-----AGCTGGCTGTAGCCCTGAGCCCTGCTCCCTCAGCC	556
DB	244	ProAlaGlyGluLeuTyrglyTpaAlaValGluThrThrProSerProGlyProGlnPro	263
QY	557	CTGCGGAGCCCGACTTGGACAATCCCACTCCC-----TTCCCAAC	598
DB	264	AlaAlaLeuThrThrGlyGluAlaAlaAlaProGluSerProHisGlnAlaGluProTyr	283
QY	599	CTGGGGCCCTCTCAGAACCCCAAGCGCTGTGTGGCGGGGAAGAGTGGAGTTC	658
DB	284	LeuSerProSerProSerAlaCysThrAlaValGlnGluPro---SerProGlyAlaLeu	302
QY	659	GAGGTGACAGCCTTCTACCGGGCGCCCAAGTCTTCAGCAGACCACTCTCGTCCCGGAG	718
DB	303	AspValThrIleMetTyrLysGlyArgThrValLeuGlnLysValValGlyHisProSer	322
QY	719	GGCTCGGCTGGTGGGG---TCCGAAGTGGGACAGACGCTGCCTGGATGCCAGTC	775
DB	323	CysThrPheLeuTyrglyProProAspProAlaValArgAlaThrAspProGlnGlnVal	342
QY	776	ACACTGCGCAGACCCCTGGCATCTCCCTGACAGAGGGAGTGATGAGTACGTGAGGCAT	835
DB	343	AlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrThrGluGlu	361
QY	836	GTGCTGAGCTCGCTGGTGGGGAGTGGCTCTCTGGCGGGCGGAGTGGCTCTGGGCC	895
DB	362	LeuLeuArgHisValAlaProGlyLeuHisLeuGluLeuLeuArgGlyProGlnLeuTrpAla	381
QY	896	CAGCGGTGGGGCACTGCACACATCTGGGCAAGTGGAGGAGTGTCTCCCAACAGC	955
DB	382	ArgArgMetGlyLysCysLysValTyrTrpGluValGly-----	394
QY	956	GGCATGGGCTGATGGCGAG-----GTCCCAAGGAC	988
DB	395	-----GlyProProGlySerAlaSerProSerThrProAlaCysLeuLeuProArgAsn	412
QY	989	AAGNAGGAGCGGCTTTGACTGGGGCCCTTCATTGTAGATCTGATTACCTTCACGGAA	1048
DB	413	CysAspThrProIlePheAspPheArgValPhePheArgGluLeuValGluPheArgAla	432
QY	1049	GGAAAGCGGACGC---TCACCAAGCTATGCCCTCTGTTCTGTGTGGGGAGTCAATGCC	1105
DB	433	ArgGlnArgArgGlySerProArgTyrThrIleTyrLeuGlyPheGlyLeuAspLeuSer	452
QY	1106	CAGNACAGCCGTGACCAAGAGGCTCGTGATGGTCAAGTGTGTGCCCACTGCCCTCAGG	1165
DB	453	AlaGlyArgProLysGluLysSerLeuValLeuValLysLeuGluProTrpLeuCysArg	472
QY	1166	GCCTGTTAGAAATGCCCGGTAGGGGTGCTCTCCTCTGGAG---AATACGTGGAC	1222
DB	473	ValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerLeuSer	492
QY	1223	CTGCATTTTCCAAAGCCACCCACTCTCCCTCACTCCGACCCAGTACAGGCGCTACCTG	1282

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 15:10:59 ; Search time 45.5 Seconds
(without alignments)
6030.994 Million cell updates/sec

Title: US-09-975-253-1
Perfect score: 2699
Sequence: 1 sggttcagtcgcccagcgc.....aaaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p_model -DEV=xl
-Q/cgn2_1/USPTO.spool_p/US09975253/runat 28102004.111510.19635/app query.fasta_1.1607
-DB=pir 79 -OPMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pc1 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09975253@cgn_1.1.77@runat 28102004.111510.19635 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	499.5	18.5	504	2 G02474	interferon regulat
2	457.5	17.0	491	2 S56753	interferon regulat
3	438.5	16.2	459	2 JC6520	interferon regulat
4	388	14.4	450	2 S57837	lymphoid-specific
5	351	13.0	425	2 A45064	interferon consens
6	343.5	12.7	424	2 A35861	transcription fact
7	301.5	11.2	393	2 A45017	lymphoid-specific
8	276.5	10.2	203	2 S57836	interferon regulat
9	240	8.9	325	2 I52998	transcription fact
10	235.5	8.7	399	2 JC4592	interferon regulat
11	232	8.6	325	2 B31595	hypothetical prote
12	226.5	8.5	1585	2 T31611	interferon regulat
13	219	8.1	349	2 A53340	interferon regulat
14	215	8.0	674	2 S13301	collagen alpha 1(X

15	214.5	7.9	744	2 S15435	collagen alpha 1(V
16	212	7.9	328	2 A36330	interferon regulat
17	211	7.8	329	2 A31595	interferon regulat
18	211	7.9	1414	1 S23809	collagen alpha 2(I
19	211	7.9	1464	2 S59856	collagen alpha 1(I
20	210	7.8	680	2 S31216	collagen alpha 1(X
21	208.5	7.8	1315	2 A56101	collagen alpha 1(X
22	208.5	7.8	1774	2 B56101	collagen alpha 1(X
23	207	7.7	1464	1 CGHUI1	collagen alpha 1(I
24	205.5	7.7	1806	1 CGHUI1	collagen alpha 1(X
25	205	7.7	680	1 CGHUI1	collagen alpha 1(X
26	204.5	7.6	1453	2 S21626	collagen alpha 1(I
27	204	7.6	633	2 B40983	collagen alpha 1(X
28	204	7.6	888	2 S28791	collagen alpha 1(X
29	203.5	7.6	1466	1 CGHU7L	collagen alpha 1(I
30	203	7.6	673	1 CGBO6C	collagen alpha 1(I
31	202.5	7.6	1603	2 S23810	collagen alpha 1(X
32	202	7.5	1418	2 T45467	collagen alpha 1(I
33	201.5	7.5	744	2 S15435	collagen alpha 1(V
34	201	7.5	438	2 S53787	collagen alpha 1(V
35	201	7.5	886	2 I50694	collagen alpha 1(I
36	201	7.5	1419	2 A41182	collagen alpha 1(I
37	201	7.5	1487	2 B41182	collagen alpha 1(I
38	200.5	7.5	744	1 A34246	collagen alpha 1(V
39	200.5	7.5	744	1 S23298	collagen alpha 1(V
40	200.5	7.5	1049	1 CGBO7S	collagen alpha 1(I
41	200.5	7.5	1460	1 EDBE1F	immediate-early pr
42	199.5	7.4	1366	1 CGHU2S	collagen alpha 2(I
43	199.5	7.5	1433	2 A46053	bullous pemphigoid
44	199	7.4	635	2 A57131	collagen alpha 2(V
45	199	7.4	743	1 S23779	collagen alpha 1(V

ALIGNMENTS

RESULT 1

G02474
interferon regulatory factor 5 - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
A:Accession: G02474
R:Grossman, A.; Mitrucker, H.W.; Lantonio, L.; Mak, T.W.
submitted to the EMBL Data Library, March 1996
A:Reference number: H01338
A:Accession: G02474
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-504 <GRO>
A:Cross-references: UNIPROT:Q13568; EMBL:U51127; NID:g1255254; PIDN:AAA96056.1; PID:g125

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Genetics:

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

A:Gene: Humirf5

C:Superfamily: lymphoid-specific interferon regulatory factor

Db 35 TrpValaAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly 54
 QY 182 GCACAGCAGGAG--GATTTCGGAATCTTCAGGCTGGCCGAGCCAGCCACTGGTCATAT 238
 Db 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTyr 74
 QY 239 GTTCCCGGAGGAGTAAGACAGCAGCTGCCAACCTCGAAGAGAAATTCGGCTCTCCCTC 298
 Db 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94
 QY 299 AACCCCAAGAGGTTGGTTTACAGAGGAGCCGAGCAAGAC--CCTCAGCACCA 355
 Db 95 AsnLysSerArgAspPheArgLeuIleLysArgGlyProArgAspMetProGlnPro 114
 QY 356 CATAAATCTACAGTTTGTGAATCAGGAGTTGGGACTTTTCCAGCCAGACACTCT 415
 Db 115 TyrLysIleTyrGluValCysSerAsnGlyPro--AlaProThrAspSerGln 131
 QY 416 CCG-----GACACCAATGGTGAGCAGTACTTCTGTATACCCAGGAACATTCGGAT 469
 Db 132 ProProGluAspTyrSerPheGlyAlaGlyGluGluGluGluGluGluGluGln 151
 QY 470 GAGTTACTGGTAACATGGTTGGCCCTCCAGTCCGAGTCCGGA----- 514
 Db 152 ArgMetLeuProSerLeuSerLeuThrAspAlaValGlnSerGlyProHisMetThrPro 171
 QY 514 ----- 514
 Db 172 TyrSerLeuLeuLysGluAspValLysTyrProProThrLeuGlnProProThrLeuGln 191
 QY 515 CCCCCAAGCTGCTGTAGCCCTGAGCCCTGAGCCCTCAGCCCTGGGAGCCGCCACTTG 574
 Db 192 ProProValLeuLeuGlyProProAlaProAspProSerProLeu--AlaProProPro 210
 QY 575 GACAATCCCACTCCCTCCCAACCTGGGCGCTCTGAG-----AACCCACTG 622
 Db 211 GlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeuGluProGlyProLeu 229
 QY 623 AAGCGGTGTGTGTCGCG--GGGAAGAG----- 649
 Db 230 ProAlaSerLeuProProAlaGlyGluGlnLeuLeuProAspLeuLeuSerProHis 249
 QY 650 -----TGGAGATTCGAGTGACAGCCTTCTACCGGCGCCGAGCTCTCCAG 697
 Db 250 MetLeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGlyArgProProArgAla 269
 QY 698 CAGACCATCTCTCCCGCGGCGCTCGCTGTGGGTCCGAAGT----- 745
 Db 270 LeuThrIleSerAsnProHisGlyCysArgLeuPheTyrSerGlnLeuGluAlaThrGln 289
 QY 746 -----GGAGACAGACGCTGCTGGATGGCGAGTCACACTGCCAGAC 787
 Db 290 GluGlnValGluLeuPheGlyProIleSerLeuGlu-----GlnValArgPheProSer 307
 QY 788 CTGGCATGCTCCCTACAGACAGGAGTGTATGATGATGATGATGATGATGATGATGATG 847
 Db 308 ProGluAspIleProSerAspLys---GlnArgPheTyrThrAsnGlnLeuLeuAspVal 326
 QY 848 CTGGGTGGGAGTGGCTCTCTGCGCGCGGCGAGTGGCTCTGGGCGCCAGCGCTGGGG 907
 Db 327 LeuAspArgGlyLeuIleLeuGlnLeuGlnGlyGlnAspLeuTyrAlaIleArgLeuCys 346
 QY 908 CACTGCCACATACATG-----GCATGAGCGAGGAGTGTCTCCCAACAGC 955
 Db 347 GlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAspSerCysProAsn--- 365
 QY 956 GGGCATGGCTGTGGCGAGTCCCAAGGACAGGAGGAGCGGTGTGGTGTGGTGTGGT 1015
 Db 366 -----ProIleGlnArgGluValLysThrLysLeuPheSerLeuGlu 379
 QY 1016 CCCTTTCATTGTATCTGATTACCTTACCGAAGGA---AGCGGACGCTACCAACGCTAT 1072
 Db 380 HisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsnThrProProPhe 399

QY 1073 GCCCTCTGTTCTGTGGGAGTCAATGGCCAGGACCATGGCCAGGAGGCTC 1132
 Db 400 GluIlePhePheCysPheGlyGluTrpProAspArgLysProArgLysLysLeu 419
 QY 1133 GTGATGGTCAAGGTTGTGCTCCACAGCTGCTCAGGGCTTGGTAGAAATGCGCCGGTAGGG 1192
 Db 420 IleThrValGlnValProValAlaAlaArgLeuLeuGluMetPheSer---Gly 438
 QY 1193 GGTGCTCTCTCCTCGGAGAACTGTGACCTGACCTGCACATTCACAGACCA----- 1242
 Db 439 GluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn--ProAspLeuLysA 458
 QY 1243 -----CCCACTCTCCCTCAC-----CTCC 1261
 Db 458 spArgMetValGluGlnPheLysGluLeuHisIleTrpGlnSerGlnGlnArgLeuG 478
 QY 1262 GACCACTACAGGCTTACCTGAGGACTTGTGGAGGCGATGATTTCCAGGCGCTGG 1320
 Db 478 InProValAlaGlnAlaProProGlyAlaGlyLeuGlyValGly---GlnGlyProTrp 496
 RESULT 2
 S56753
 Interferon regulatory factor 3 - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S56753
 R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
 Nucleic Acids Res. 23, 2137-2146, 1995
 A:Title: cIRF-3, a new member of the interferon regulatory factor (IRF) family that is
 A:Reference number: S56753; MUID:95334365; PMID:7541908
 A:Accession: S56753
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-491 <R>
 A:Cross-references: UNIPROT:Q90643; EMBL:U20338; NID:g790580; PID:AAA86995.1; PID:g7905
 C:Superfamily: lymphoid-specific interferon regulatory factor
 Alignment Scores:
 Pred. No.: 1,18e-21 Length: 491
 Score: 457.50 Matches: 146
 Percent Similarity: 41.57% Conservative: 61
 Best Local Similarity: 29.32% Mismatches: 174
 Query Match: 16.95% Indels: 117
 DB: 2 Gaps: 19
 US-09-975-253-1 (1-1426) x S56753 (1-491)
 QY 59 AAGCCACGGNTCCCTGCTGGTGTGTCAGCTGACCTGGGCAACTGGAGGCGTG 118
 Db 12 LysLeuArgPheGlyProTrpLeuLeuAsnAlaValSerSerGlyLeuTyrArgGlyLeu 31
 QY 119 GCCTGGGTGAACAGAGCCGACGCGCTTCCGATCCCTTGGAGCAGCGCTTACGCGCAG 178
 Db 32 CysTrpIleAspProAspArgArgIlePheArgIleProTrpLysHisAsnAlaArgLys 51
 QY 179 GATGCACAGCAGGAGGATTTCCGAATCTCCAGCCTGGCGCCAGGCACTGGTGCATAT 238
 Db 52 AspValThrSerSerAspValGluIlePheLysAlaTrpAlaLysAlaSerGlyArgTyr 71
 QY 239 GTTCCCGGAGGATTAAGCCAGACCTGCCAACCTGGAAGAGGAATTTCCCTCTGCCCTC 298
 Db 72 GluGlyAsnAlaGluAspPro-----AlaLysTrpLysThrAsnPheArgCysAlaLeu 89
 QY 299 AACCGCAAGAGGTTGCGTTTAGCAGAGGACCGGAGCAGGACCTCAGACCCACAT 358
 Db 90 ArgSerThrHisMetPheMetLeuLeuGluAspArgSerLysCysAsnAspAspProHis 109
 QY 359 AAAATCTACAGTTTGTGAACCTCAGAGTT----- 388
 Db 110 LysValTyrAla---ValAlaSerGlyValProAsnAspArgGlySerGlyProVal 128
 QY 389 -----GGGACTTTTCCAGCCA 406

Fri Oct 29 11:11:44 2004

```

Db      315 AspGluValValGlnValPheAspThrSerGlnPhePheArgLeuLeuGlnGlnPheTyr 334
      |||::: ||||| ||| :||| |||
QY      1046 GAAGGACGGACGCTACCACGCTATGCCCTGCTGCTGTGTGGGGAGTCAATGGCCC 1105
      ||||| ||| :||| ||| :|||
Db      335 AsnSerGlnGlyArgLeuProAspGlyArgValValLeuCysPheGlyGluGluPhePro 354
      ||||| ||| :||| ||| :|||
QY      1106 CAGGACCACGCCGTGACCAAGAGGCTGCTGATGGTCAAGGTTGTGCCACGTGCTCAGG 1165
      ||| :||| :||| :||| :|||
Db      355 AspMetAlaProLeuArgSerIysLeuIleLeuValGlnIleGluGlnLeuTyrValArg 374
      ||| :||| :||| :||| :|||
QY      1166 GCCTTGGTAGAAATGCC---CGGTAGGGGGTGCCCTCTCCCTG 1207
      ||| ||| ||| :||| :|||
Db      375 GlnLeuAlaGluGluAlaGlyLysSerCysGlyAlaGlySerVal 389
      ||| ||| ||| :||| :|||

RESULT 6
A35861
interferon consensus sequence-binding protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 09-Jul-2004
C/Accession: A35861
R/Driggers, P.H.; Ennlist, D.L.; Gleason, S.L.; Mak, W.H.; Marks, M.S.; Levi, B.Z.; Flana
Proc. Natl. Acad. Sci. U.S.A. 87, 3743-3747, 1990
A/Title: An interferon gamma-regulated protein that binds the interferon-inducible enhan
A/Reference number: A35861; MUID:90251633; PMID:2111015
A/Accession: A35861
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-424 <DRI>
A/Cross-references: UNIPROT:P23611; GB:M32489; NID:gl94088; PIDN:AAA37878.1; PID:G3093266
C/Superfamily: lymphoid-specific interferon regulatory factor
C/Keywords: DNA binding; transcription regulation

```

Alignment Scores:	2.25e-14	Length:	424
Pred. No.:	343.50	Matches:	117
Score:		Conservative:	54
Percent Similarity:	41.71%	Mismatches:	180
Best Local Similarity:	28.54%	Indels:	59
Query Match:	12.73%	Gaps:	14
DP:	2		

TS-09-975-253-1 (1-1426) x A35861 (1-424)

65	QY	CGGNTCTCGCCCTGGCTGGTTCGCAGCTGCAGCACTGGGGCAACTGGAGGCGGTGGCTGG	124
	Db		
9	ArgLeuArgGlnTrpLeuIleGluGlnIleAspSerMetTyrProGlyLeuIleTrp	28	
125	QY	GTGAACAAGACCGCAGCGGCTTCGCGATCCCTTCGAAAGACGCGCTACGGCAGGATGCA	184
	Db		
29	GluAsnAspGluLysThrMetPheArgIleProTrpLysHisAlaGlyLysGlnAspTyr	48	
185	QY	CAGCAGAG---GATTTTCGGATCTTCAGGCGCTGGCGGAGGCCACTGGTCATATGTT	241
	Db		
49	AsnGlnGluValAspAlaSerIlePheLysAlaTrpAlaValPheLysGlyLysPheLys	68	
242	QY	CCGGGAGGGATAAAGCCAGACCTGCCAACCTGGAAAGAGGAATTTCCGCTCTGCCCTCAAC	301
	Db		
69	GluGly---AspLysAlaGluProAlaThrTrpLysThrArgLeuArgCysAlaLeuAsn	87	
302	QY	CGCAAAAGAAGGTTTCGTTTACGACAGACCGGACGAAAG---GACCCTCAGACCCACAT	358
	Db		
88	LysSerProAspPheGluGluValThrAspArgSerGlnLeuAspIleSerGluProTyr	107	
359	QY	AAATCTACGAGTTTGTG-----AACTCAGAGTTGGGACTTT	397
	Db		
108	LysValTyrArgIleValProGluGluGlnLysCysLysLeuGlyValAlaProAla	127	
398	QY	TCCAGCCAGACACCTCTCCGACACCAATGGTGGAGGAGTACTTCTGAT-----ACC	451
128	GlyCysMetSerGluValProGluMetGluCysGlyArgSerGluIleGluGluLeuIle	147	
452	QY	CAGGAAGACATTTCTGGATGAGTTTACTGGGTAAACATGGTGTGGCCCCCACTCCCGAGATCCG	511

148 LysGluProSerValAspGluTyrMetGlyMetThrLysArgSerProSerPro----- 165
512 GGRACCCCHAGCCTGGCTGTACCCCTCGAGCCCTGCCCTCAGCCGCTG----- 559
166 -----ProGluAlaCysArgSerGlnIleuLeuProAspTrp 177
560 -----CGAGAGCCCGAGCTTGACAAATCCC-----ACTCCCTTCCCAACCTGGGGCCC 607
178 TrpValGlnGlnProSerAlaGlyLeuProLeuValThrGlyTyrAlaAlaTyrAspThr 197
608 TCTGAGAACCCACTGAAGCGGCTGTGGTCCGGGGGAGAGTGGGAGTTTCAGAGTGACA 667
198 HisHisSerAlaPheSerGlnMetVal-----IleSer 208
668 GCCTTTACCGGGCCCGCAAGTCTTCAGACAGACCATCTCTGCCGCGAGGCGCTGGGG 727
209 PheTyrTyrGlyGlyLeuValGlyGlnAlaThrThrCysLeuGluGlyCysArg 228
728 CTGTGGGGTCCGAAGTGGGAGACAGACGCTGCCTGGA-----TGGCCAGTCT 775
229 LeuSerLeuSerGlnProGlyLeuProLysLeuTyrGlyProAspGlyLeuGluProVal 248
776 ACATGCCACACCGCTGGCATGCCCTCCACAGACAGGGAGTGATGAGTACCTGAGGCAT 835
249 CysPheProThrAlaAspThrIleProSerGluArg---GlnArgGlnValThrArgLys 267
836 GTGCTGAGCTGCTGGTGGGGACTGGGTCTCTGGCGGGCGCGGAGTGGTCTCTGGGCC 895
268 LeuPheGlyHisLeuGluArgGlyValLeuLeuHisAsnArgLysGlyValPheVal 287
896 CAGCGGTGGGGCATTGCCACACATCTGGGCGAGTGAGCGAGGAGCTGTCCCCACAGC 955
288 LysArgLeuCysGlnGlyArgValPheCys-----Ser 298
956 GGGCAT-----GGGCTGTAGGGGAGTCCCCAAGACACAGGAAGGAGGC 1000
299 GlyAsnAlaValValCysLysGlyArgProAsmLysLeuGluArgspGluValGln 318
1001 GTGTTTGACCTGGGGCCCTTCATTGTAGATCTCATTTACCTTCACGGAAGAACGGACGC 1060
319 ValPheAspThrAsnGlnPheIleArgGluLeuGlnPheTyrAlaThrGlnSerArg 338
1061 TCACACGCTATGCCCTCTGGTCTGTGTGGGGAGTCTATGGCCCGACGAGCAGCGCTGG 1120
339 LeuProAspSerArgValValLeuCysPheGlyGluGluPheProAspThrValProLeu 358
1121 ACCAAGAGCTCGTGTATGGTCAAGTTGTGCCACGTCCTCAGGGGCTTGGTAGAAATG 1180
359 ArgSerLysLeuIleLeuValGlnValGluGlnLeuTyrAlaArgGlnLeuValGluGlu 378
1181 GCC---CGGGTAGGGGGTGCTCTCCCTGCT 1207
379 AlaGlyLysSerCysGlyAlaGlySerLeu 388

RESULT 7
A45017
transcription factor ISGF3 gamma chain - human
N:Alternate names: alpha-interferon-responsive transcription factor
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A45017
R/Reals, S.A.; Schindler, C.; Leonard, D.; Fu, X.Y.; Aebersold, R.; Darnell Jr. J.
Mol. Cell. Biol. 12, 3315-3324, 1992
A>Title: Subunit of an alpha-interferon-responsive transcription factor is rela
A/Reference number: A45017; MUID: 92334329; PMID: 1630447
A/Accession: A45017
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-393 <A>
A/Cross-references: UNIPROT: Q00978; GB: M87503; NID: g184652; PIDN: AAA58687.1; P
A/Experimental source: Hela cells
A/Note: sequence extracted from NCBI backbone (NCBI:108663)
A/Note: part of this sequence was confirmed by protein sequencing

2. **DESIGN.**

RESUL
B45017

A43017 transcription factor TSGF3 gamma chain - human

transcription factor 13013 gamma chain

N; Affiliates names: arpiia - lico
C. Species: Homo sapiens (man)

C;Species: HOMO SAPIENS (MAN)
C.Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 09-Jul-2004

C; Date: 30-Sep-1993
C: Accession: A45017

C;Accession: A45017

Author	Year	Sample Size	Prevalence (%)
R;Veals, S.A.; Schindler, C.; Leonard	1992	12	3315-3324
Wal Cell Picl	12	3315-3324	1992

Mol. Cell. Biol. 12, 3315-3324, 1992

A; Title: Subunit of an alpha-interferon
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580
 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680
 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700
 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720
 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740
 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780
 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800
 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900
 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 99

A;Reference number: A45017; MUID:923

A;Accession: A45017

A; Status: nucleic acid sequence not

A;Molecule type: mRNA

A;Residues: 1-393 <VEA>

A;Cross-references: UNIPROT:Q00978;

A; Experimental source: HeLa cells

A;Note: sequence extracted from NCBI

A;Note: part of this sequence was co

C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription factor

Alignment Scores:

Pred. No.: 1.08e-11 Length: 393
Score: 301.50 Matches: 108
Percent Similarity: 41.46% Conservative: 62
Best Local Similarity: 26.34% Mismatches: 177
Query Match: 11.17% Indels: 63
DB: 2 Gaps: 16

US-09-975-253-1 (1-1426) x A45017 (1-393)

```

QY 38 GGGCGGACCATGGGAGACCCCAAGCCACCGTCTGCTCCCTGGCTGGTGTGCTGGAGCTGGAC 97
Db 4 GlyArgAlaArgCysThrArgLysLeuArgAsn-----TrpValValGluGlnValGlu 21
QY 98 CTGGGCAACTGGAGGCGCTGGCTGGTGAACAAGAGCGGACGCGCTTCGCGATCCCT 157
Db 22 SerGlyGlnPheProGlyValCysTrpAspAspThrAlaLysThrMetPheArgIlePro 41
QY 158 TGAAGACAGCGCTACGGCAGGAT---GCACAGCAGGAGATTTCGGAATCTTCCAGGCC 214
Db 42 TrpLysHisAlaGlyLysGlnAspPheArgGluAspGlnAspAlaAlaPhePheLysAla 61
QY 215 TGGGCGGAGCCACTGCTGCATATGTCGCGGAGGAGGATAAGCCACCTGCACTGG 274
Db 62 TrpAlaIlePheLysGlyLysLysGluGly---AspThrGlyGlyProAlaValTrp 80
QY 275 AAGAGGAATTCGCTCTCCCTCAACCGCAAGAGGGTGGTGTAGCAGAGGACCGG 334
Db 81 LysThrArgLeuArgCysAlaLeuAsnLysSerSerGluPheLysGluValProGluArg 100
QY 335 AGCAAG---GACCTCAGACCCACATAAATCTACAGTTGTGTAACCTCAGAGTTGGG 391
Db 101 GlyArgMetAspValAlaGluProTyrLysValTyrGlnLeuLeuProGlyLysVal 120
QY 392 GACTTTTCCAGCCAGACACTCT-----CCGACACCAAGTGGTGGAGGAGTACTTCT 445
Db 121 Ser---GlyGlnProGlyThrGlnLysValProSerLysArgGlnHisSerValSer 139
QY 446 GATACCCAGAGACATCTCGATGAGTACTGGGTAAACATGTTGGCCCACTC--- 502
Db 140 SerGluArgLysGlu---GluGluAspAlaMetGlnAsnCysThrLeuSerProSerVal 158
QY 503 -----CGAGAT 508
Db 159 LeuGlnAspSerLeuAsnAsnGluGluGlyAlaSerGlyGlyAlaValHisSerAsp 178
QY 509 CCGGACCCCAAGCTGCTAGCCCTGAGCCCTGAGCCCTGCGCTCAGCCCTGGAGCCCC 568
Db 179 IleGlySerSerSerSerSerSerSerProGluPro-----Gln 191
QY 569 AGCTTGGACAAATCCCACTCCCTTCCCAACCTGGGCGCCCTCTGAGAACCCACTGAAGCG 628
Db 192 GluValThrAspThrThrGluAlaProPheGlnGly---AspGlnArgSerLeuGluPhe 210
QY 629 CTGTTGGTCCGGGAGAGAGTGGAGTTGAGGTTGACAGCCCTTCTACCGGCGCCGCCAA 688
Db 211 LeuLeuProGluProAspTyrSerLeuLeuLeuThrPheIleTyrAsnGlyArgVal 230
QY 689 GTCTTCCAG-----CAGACCATCTCTCCCGGAGGCGCTGGCTGGTGGGTGCC 739
Db 231 ValGlyGluAlaGlnValGlnSerLeuAspCys-----ArgLeuValAlaGlu 246
QY 740 GAAGTGGGAGACAGGACGCTGCTGGATGGCGGAGTCAACCTGCCAGACCTGCCATGTCC 799
Db 247 ProSerGlySerGluSerSerMetGlu---GlnValLeuPheProLysPro----- 262
QY 800 CTGACAGACAGGGAGTGAATGATGATGAGGCAATGCTGAGGCAATGCTGAGTGGTGGGGA 859
Db 263 -----GlyProLeuGluProThrGlnArgLeuLeuSerGlnLeuGluArgGly 278

```

```

QY 860 CTGGCTCTCTGGCGGCGCCGGCAGTGGCTCTGGGCGCCAGCGGCTGGGCGCACTGCCACACA 919
Db 279 IleLeuValAlaSerAsnProArgLysLeuPheValGlnArgLeuCysProIleProIle 298
QY 920 TACTGGCAGTGAAGCGAGGAGTGTCCCAACAGCGGCGCATGGCTGTATGCGGAG--- 976
Db 299 SerTrp-----AsnAlaProGlnAlaProGlyProGly 310
QY 977 -----GTCCCAAGACAGGAGGAGGCGTGTGGCTGTGGCTGGGCGCTTCATTGTA 1027
Db 311 ProHisLeuLeuProSerAsnGluCysValGluLeuPheArgThrAlaLysPheCysArg 330
QY 1028 GATCTGATTACCTTACGGAAGACGAGCGCTCACCGCTATGCCCTCTGCTGTCTGT 1087
Db 331 AspLeuValArgTyrPheGlnGlyLeuGlyProProLysPheGlnValThrLeuAsn 350
QY 1088 GTGGGCGAGTCATGGCCCGCAGGACCGCTGGACCAAGAGGCTGCTGATGGTCAAGTT 1147
Db 351 PheTrpGluGluSerHisLysSerHisThrProGlnAsnLeuIleThrValLysMet 370
QY 1148 GTCCCACTGCTCAGGCGCTTGGTAGAA 1177
Db 371 GluGlnAlaPheAlaArgTyrLeuLeuGlu 380
RESULT 8
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S57836
R;Matsuyama, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, F.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A;Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A;Reference number: S57836; MUID:95334364; PMID:7541907
A;Accession: S57836
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-203 <MAT>
A;Cross-references: EMBL:U11692
C;Superfamily: lymphoid-specific interferon regulatory factor
Alignment Scores:
Pred. No.: 4.18e-10 Length: 203
Score: 276.50 Matches: 77
Percent Similarity: 31.25% Conservative: 33
Best Local Similarity: 21.88% Mismatches: 79
Query Match: 10.24% Indels: 163
DB: 2 Gaps: 6
US-09-975-253-1 (1-1426) x S57836 (1-203)
QY 77 TGGCTGTGTGCGACGTGGCACTGGAGGCGCTGGCTGGTGAACAGAGC 136
Db 5 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGlu 24
QY 137 CGACGCGCTTCGCGATCCCTTGAAGCAGCGCTTACGCGAGGAT---GCACAGCAGGAG 193
Db 25 LysSerValPheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGlu 44
QY 194 GATTTCGAATCTCCAGCGCTGGCGGCGGAGCCACTGTGTCATATGTTCCCGGAGGAT 253
Db 45 AspAlaAlaPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeu 64
QY 254 AAGCCAGACCTGCCAACCTGGAAGAGGAATTTCCGCTCTGCGCTCAACCGCAAGAGGG 313
Db 65 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 84
QY 314 TTGCGTTTAGCAGAGCCGCGGCAAG---GACCTCAGCAGCCCACTAAATCTTACGAG 370
Db 85 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg 104
QY 371 TTTGTGAACCTCAGGAGTTGGGACTTTTCCAGCGCAGACACTCTCCGGAACCAATGGT 430

```

Fri Oct 29 11:11:44 2004

us-09-975-253-1.1.n2p.rpx

Score: 240.00 Matches: 65
Percent Similarity: 43.66% Conservativity: 28
Best Local Similarity: 30.52% Mismatches: 85
Query Match: 8.89% Indels: 35
DB: 2 Gaps: 6

US-09-975-253-1 (1-1426) x 152998 (1-325)

QY 47 ATGGAGACCCCAAGCCACGGTCTCGCTGGCTGGTGTCTGACGTGGACCTGGGGCAA 106
DB 1 MetProLeuThrArgMetArgMetArgProThrLeuGluMetGlnLeuMetSerAsnGln 20
QY 107 CTGAGGCGCTGGCTGGCTGAACAGAGCCGACCGCTTCCGATCCCTTGGAGACAC 166
DB 21 IleProGlyLeuIleThrPheLeuMetIlePheGlnIleProThrLysHis 40
QY 167 GGCCTACGCGAGGATGACAG---CAGGAGGATTTCCGAATCTCCAGGCTGGGCGGAG 223
DB 41 AlaAlaLysHisGlyTrpAspIleAsnLysAspAlaCysLeuPheArgSerTrpAlaIle 60
QY 224 GCCACTGGTGATATGTTCCCGGAGGATAGCCAGACCTGCCAACCTGGAGAGAAAT 283
DB 61 HisThrGlyArgTrpLysAlaGlyGluLysGluProAspProLysThrTrpLysAlaAsn 80
QY 284 TTCCGCTCTGCTCAACCGCAAGAGGGTTCGTTAGCAGAGACCGGACCAAGGAC 343
DB 81 PheArgCysAlaMetAsnSerLeuProAspIleGluGluValLysAspGlnSerArgAsn 100
QY 344 CCT---CAGACCCACATAAATCTACGAGTTGTG----- 376
DB 101 LysGlySerSerAlaValArgValTyrArgMetLeuProLeuThrLysAsnGlnArg 120
QY 377 -----AACTCAGGAGTTGGGACCTTTTCCCAG 403
DB 121 LysGluArgLysSerLysSerSerArgAspAlaLysSerLysAlaLysArgLysSerCys 140
QY 404 CCAGACACCTCTCCGACACCAATGTGGAGGAGTACTTCTGATACCCAGGAGACATT 463
DB 141 GlyAspSerSerProAspThrPheSerAspGlyLeuSerSerSerThrLeuProAsp-As 160
QY 464 CTGGATGAGTACTGGTAAACATGGTTCGGCCCACTCCAGATCCGACCCCAAGC 523
DB 160 PHisSerSerThrValProGlyTyrMetGlnAspLeuGluValGluGln----- 177
QY 524 CTGCTGTAGCCCTCGAGCCCTCGCTCAGCCCTCGGAGGCC-----CCAGC 571
DB 178 -----AlaLeuThrProAlaLeuSerProCysAlaValSerSerThrLeuProAs 194
QY 572 TTGGACAAATCCCA-----CTCCCTTCCCAACC 599
DB 194 pTrpHisIleProValGluValValProAspSerThr 206

RESULT 10

JC4592
transcription factor ISGF3 gamma chain - mouse
N:Alternate names: interferon-stimulated gene factor 3 gamma chain
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C:Accession: J04592; S71599
R:Suhrara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
J. Biochem. 119, 231-234, 1996
A:Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)
A:Reference number: J04592; MUID:97037063; PMID:8882710
A:Accession: J04592
A:Molecule type: mRNA
A:Residues: 1-399 <SUH>
A:Cross-references: UNIPROT:Q61179; EMBL:U51992; NID:G1263309; PIDN:AA52494.1; PID:G1263309
A:Experimental source: L929 cells
R:Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.
FEBS Lett. 358, 225-229, 1995
A:Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced e
A:Reference number: S71599; MUID:95145714; PMID:7843405
A:Accession: S71599

105 IleVal----- 106
QY 431 GGAGCAGTACTTCTGATACCAGGAGACATTTCTGGATGAGTTACTGGTAACATGTGT 490
DB 106 ----- 106
QY 491 TTGGCCCCACTCCAGATCCGGGACCCCAAGCCTGGCTGTAGCCCTGAGCCCTGGCCCT 550
DB 106 ----- 106
QY 551 CAGCCCCCTGGGAGCCCGAGCTTGGACATCCCACTCCCTTCCCAACCTGGGGCCCTCT 610
DB 106 ----- 106
QY 611 GAGAACCCACTGAAGCGGCTGTTGGTCCCGGGGGAAGTGGGAGTTGAGGTGACAGCC 670
DB 106 ----- 106
QY 671 TTCTACCGGGCGGCAAGTCTTCAGAGACCATCTCTGCGGAGGCGCTGGCGCTG 730
DB 107 -----ProGluGlyGlnArg--- 111
QY 731 GTGGGTCCGAAGTGGGACAGACAGCGCTGCTGGATGCCAGTCACTGCCAGACCCT 790
DB 111 ----- 111
QY 791 GGCATGTCCCTGACAGACAGAGGAGTGTAGTACGTAGGCGATGTGTGAGTGTGCTGCT 850
DB 112 -----LysAsnIleGluLysLeuLeuSerHisLeu 121
QY 851 GGTGGGGGACTGCTCTCTGGCGGGCGGCGAGTGTGCTGAGGCGATGTGTGAGTGTGCTG 850
DB 122 GluArgGlyLeuValLeuTrpMetAlaProAspGlyLeuTyrAlaLysArgLysGln 141
QY 911 TGCCACATACTGGGCGAGTGAGCGAGGAGTGTCTCCCAACAGCGGGCATGGGCGCT--- 967
DB 142 SerArgIleTyrTrp-----AspGlyProLeu 150
QY 968 -----GATGGCGAGTCCCAAGGACCAAGGAGGCGGTGTGAC 1009
DB 151 AlaLeuCysSerAspArgProAsnLysLeuGluArgAspGlnThrCysLysLeuPheAsp 170
QY 1010 CTGGGGCCCTTCACTGTAGTACTGATCTTACCTTACGGAAGGAGCGACCTCACACGC 1069
DB 171 ThrGlnGlnPheLeuSerGluLeuGlnValPheAlaHisGlyArgProAlaProArg 190
QY 1070 TATGCCCTCTGTTCTGTGGGGGAGTCAATGGCCC 1105
DB 191 PheGlnValThrLeuCysPheGlyGluGluPhePro 202

RESULT 9

152998
interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: J52998
R:Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deisseroth, A.B.
DNA Cell Biol. 11, 605-611, 1992
A:Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: J52998; MUID:93000481; PMID:1382447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: UNIPROT:P10914; GB:L05072; NID:G184648; PIDN:AAA36043.1; PID:G184649
C:Genetics: GDB:IRF1
A:Gene: GDB:IRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: 5q31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1
Alignment Scores: 9.13e-08 Length: 325
Pred. No.:

THE UNIVERSITY OF CHICAGO

US-09-975-253-1 (1-1426) x B31595 (1-325)

1384	QY	GGTTGAGTGTGGGAAACAGGGGGTTGGAGGCACCATGAGAGCGAGGGCTCAGCT	1322
1010	Db	lyGlyGlyGlyGlySerSerGlyGlyTyr	1024
1324	QY	CTCCCGAGGGCCCTGGAAATCATCCCTCCACCAAGTCTCTCAGGTAGGCGCTTGACT	1266
1025	Db	AlaPro--ProProProProProProProProProProAlaProAlaProAlaProS	1044
1265	QY	-----GGTCGGAGGTGAGGGAGAGTGGTGCTGTGGMAATGTG	1226
1044	Db	erSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyGlyS	1064
1225	QY	CAGGTCCACAGTATTCTCCAGGAGGAGGACACCCCTACCCGGGCCATTTCTACCAAGGC	1166
1064	Db	erSerGlyGlyTyrSerGlyGlySerThrAlaProProProProProProProProp	1084
1165	QY	CTGAGGACGCTGGGCACAACTTGACCATCAGGACCTCT	1121
1084	Db	roProProAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSerGlyGlyS	1104
1120	QY	CAAGGCTGGTCTGGGGCCATGACTCCCCCACACAGAACCAGAGGCGCATGCGTGTA	1061
1104	Db	erSerGlyGlySerAlaAlaGlyGlyGlyGlyGlySerSerGlyGlyTyrSerGlyGlyS	1124
1060	QY	GGCTCCGCTTCCTTCGCTGAAGTAATCATGATCTACATGAAGGGCCCCAGGTCAACAC	1001
1124	Db	erAlaAlaProProPro-	1129
1000	QY	GCCTCTCTCTTCTCTGGGGACCTCGCCATCAGGCCCATGCCGCTGTGGGAGCAG	941
1130	Db	-----ProProProProProProProProAlaProAlaProAlaProSerSerG	1147
940	QY	CTCCTCGCTCACTGCCCGCATGTGTGGCAGTGCCTCCAGCGCGCTGGG	894
1147	Db	lyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyGlySerSerG	1167
893	QY	-----CCCAGAGCCACTGCCCGGC	875
1167	Db	lyGlyTyrSerGlyGlySerAlaAlaProProProProProProProProProProp	1187
874	QY	CGCCAGAGCCAGTCCCCCAGCCAGCTCAGCACATGCTCAGTAGCTCATCAC	815
1187	Db	roAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSerGlyGlySerSer	1205
814	QY	TCCCTCTCTCTCAGGACATGCCA-----GGGTCTGGCAGTGTGACTGCCCATCCAG	762
1206	Db	-----GlyGlySerAlaAlaGlyGlyGlyGlyGlySerSerGlyGlyTyrSerG	1222
761	QY	GCAGCTCTCTCTCCCATTCGACCCACCAGCCGCGCAGCCCTCCG	714
1222	Db	ly-GlySerAlaAlaProProProProProProProProProProAlaProAla	1241
713	QY	-----GGCAGGAGATGGTCTGCTGGAAGACTTGGCGGC	681
1242	Db	ProAlaProAlaProSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAla	1266
680	QY	-----CCCGTAGAGGCTGTCACTCGAACTCCC	651
1262	Db	GlyGlyGlyGlyGlySerSerGlyGlyTyrThrGlyGlySerAlaAlaProProPro	1281
650	QY	ACTCTTCCCGGCACCAACAGCCGCTTCAGTGGGTCTCAGAGGGCCCCAGGTTGGGA	591
1282	Db	ProProProProProProProProProProAlaProAlaProAlaProSerSerGly	1301
590	QY	AGGGAGTGGGATGTCGAAGCTGGGGCTCCGCA-----GGGGCTGAGGGCAGGCTCAG	537
1302	Db	GlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyGlySerSerGly	1321
536	QY	GGGCTACAGCCAGGCTTGGGGTCCGGATCTGGGAGTGGGGCCCAACCATGTTACCCA	477
1322	Db	GlyTyrSerGlyGlySerAlaAlaProProProProProProProProProProPro	1341

776 ACAGTCCAGACCTGGCATGCTCCCTGACA 805
 282 ThrSerAenLysProAspLeuGlnValThr 291

RESULT 14
 S13301
 collagen alpha 1(X) chain precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
 C:Accession: S13301
 R:Thomas, J. T.; Kwan, A. P. L.; Grant, M. E.; Boot-Handford, R. P.
 Biochem. J. 273, 141-148, 1991
 A:Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen. EV
 A:Reference number: S13301; MUID:91113131; PMID:1703407
 A:Accession: S13301
 A:Molecule type: mRNA
 A:Residues: 1-674 <THO>
 A:Cross-references: UNIPROT:P23206; EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264
 C:Genetics:
 C:Gene: COL10A1
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
 F:547-673/Domain: complement C1q carboxyl-terminal homology <CIQ>

Alignment Scores:
 Pred. No.: 3,7e-06 Length: 674
 Score: 215.00 Matches: 139
 Percent Similarity: 34.06% Conservative: 33
 Best Local Similarity: 27.52% Mismatches: 184
 Query Match: 8.03% Indels: 149
 DB: 2 Gaps: 31

US-09-975-253-1 (1-1426) x S13301 (1-674)

QY 1340 GAGCGAGGCTCAGCTCTCCAGGCGCCCTGGAATCCA-----TGCCCTCCACCAAGT 1287
 DB 58 GluGlnGlyIleProGlyProGlyProGlyProAlaGlyProArgGlyHisProGlyProSer 77
 QY 1286 ---CCTGCAGTAGGCTTGTACTGTCGCGAGGTGAGGAGAGTGGGTGCTGTTGGAAA 1230
 DB 78 GlyProProGlyLysProGlyThrGly----- 86
 QY 1229 TGTGCAGGTCCACAGTATTTCTCAGGAGGAGGACCCCTACCGGGCCATTTCTACCA 1170
 DB 87 SerProGlyProGlnGlyGlnProGly-----LeuProGlyPro-----Pro 100
 QY 1169 AGGCCCTGAGCA-----CG 1155
 DB 101 -GlyProSerAlaThrGlyLysProGlyLeuProGlyLeuProGlyLysGlnGlyGluAr 120
 QY 1154 TGGGCACACCTTGACCATCAGAGCCTTTGTGTCACCGGTGTCTCGGGGCA---TG 1098
 DB 120 gGlyLeuasn-----GlyProLysGlyAspIleGlyProAlaGly 133
 QY 1097 ACTCCCGCCACACAGAA-----CCAGAGGGCATACCGGTGTGAGCGTCCCGCTTC 1050
 DB 133 yLeuProGlyProArgGlyProProGlyProGlyIleProGlyProAlaGlyIleSe 153
 QY 1049 CTTCC-----CGTAGAGTATACATCTACAATGAAGGCCCCAGCTCAACACAGCCTC 996
 DB 153 rValProGlyLysProGlyProGln-----GlyProThrGlyGluProGlyPr 169
 QY 995 CTTCTGTGCTTGGGACCTCGCCATCAGGCCCATGCCCTGTGTGGGAGCAGCTC-- 938
 DB 169 oArgGlyPheProGlyGluLysGlyThrSerGlyValProGlyLeuasnGlyGlnLysG 189
 QY 937 ----CTCGCTCACTGCCAGTATGT-----GTGGCATGTGCCCGAGCGCTGGG 894
 DB 189 yGluMetGlyHisCysThrProCysArgProGlyGluArgGlyLeuProGlyProGlnG 209
 QY 893 CCAGAGCACCATTGCCCGCGCGGCA-----GAGAGCCAGTCCCCCAC 852

209 yProThrGlyProProGlyProProGlyValGlyLysArgGlyGluasnGlyLeuProG 229
 851 CCAGGCAGCTCAGCACATGCTCCTCAGTAGCTCATCATCTCCCTCTGTCTCAGGACATGC 792
 229 yGln-----ProGlyLeuLysGlyAsp-----GlnGly-ValP 240
 791 CAGGGTCTGGCAGTGTGAGTGGCCATCCAGGAGGCTCTGTCTCCCATCTCGGACCCCA 732
 240 roGlyGluArgGlyAlaAlaGlyProSerGly-----ProGlnGlyProProG 256
 731 CCAGCCGAGGCCCT-----CCGSGCAGGAGATGCTCTGCTGGAAGACTT 687
 256 lyGluGln-GlyProGluGlyIleGlyLysProGlyAlaProGlyIleProGlyGlnPro 275
 686 GSGCGCCCGGTAGAGGCTGTCACTCGAACTCCCACTCTTCCCGCCGACCCACAGACC 627
 276 GlyIleProGlyMetLysGlyGlnProGlyAlaProGlyThrAlaGlyLeuProGlyAla 295
 626 GCTTACGTGGTCTCAGAGGCCCCAGTTTG-----GGAAGGAGTGGATTGTCCAAGC 570
 296 ProGlyPheGlyLysProGlyLeuProGlyLysGlyGlnArgGlyProValGlyLeu 315
 569 TGGGCTCCGAGGGCTCAGGAGGCTCAGGGCTCAGGGCTA-----CAGCCA 525
 316 ProGlySerProGlyAlaLysGlyGlu---GlnGlyProAlaGlyHisProGlyGluAla 334
 524 GGCTTGGGGTCCCGATCTGGAGTGGGGCCCAACACCATGTTTACCCAGTAACATCATCA 465
 335 GlyLeuProGlyProSerGlyAsnMetGlyPro-----GlnGlyPro 348
 464 GAATGCTTCTCGGTATCAGAAGTACTGCCTCCACCATGTTGTCCGAGAGGTGTCTG 405
 349 LysGlyIleProGly----- 353
 404 GCTGGGAAAGTCCCACTCTCAGTTCACAACTCGTAGATTTTATGTGGTCTGTGAG 345
 354 -----AsnProGlyLeuProGlyProLysGlyGlu-----MetGlyProVal 367
 344 GGTCTTGTCTCCGTCTCTCTGCTAAACGCAACCTCTTTTGGGTTGAGGGCAGACGGA 285
 368 GlyProAlaGlyAsnPro-----GlyAlaLysGly 377
 284 ---AATCTCTCTCCAGTTGGCAGGTCTG---GCTTATCCCTCCCGGAAACATATGCAC 231
 378 GluArgGlySerSerGlyLeuAspGlyLysProGlyTyProGlyGluProGlyLeuAsn 397
 230 -----CAGTGGCTCCGCGCCAGGCTCGAAGATTCCGAAAT 195
 398 GlyProLysGlyAsnProGlyLeuProGlyProGlyLysAspProGlyIleAlaGlySer 417
 194 CCTCTGTGTGCATCTCGCGTAGGCGGTCTTCAAGGGATGCGAAGCGCTCGCGC 135
 418 ProGlyLeuProGlyProValGlyProAlaGlyAlaLysGlyValProGlyHisAsnGly 437
 134 TCTTGTCCACGAG-----CCAGCCCTCCAGTCCCGCCAGGTCCAGCTCGCACACCA 81
 438 GluAlaGlyProArgGlyValProGlyIleProGlyThrArgGlyProIleGlyProPro 457
 80 GCCAGGCGAGNCCGTGCTTGGGTCCCATGCTCCGCTCATGATGGAAGGTGGG 21
 458 Gly-----IleProGlyPhePro-----GlySerLysGlyAspValGly 470
 20 GGTGCGG 12
 471 ThrProGly 473

RESULT 15
 S15435
 collagen alpha 1(VIII) chain precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S15435

Search completed: October 28, 2004, 15:43:01
Job time : 66.5 secs

This Page Blank (copy)

LENGTH: 427 amino acids	
TYPE: amino acid	
TOPOLOGY: linear	
MOLECULE TYPE: protein	
US-08-705-771-16	
Alignment Scores:	
Pred. No.:	3,82e-178
Score:	2326.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	86.18%
DB:	3
US-09-975-253-1 (1-1426) x US-08-705-771-16 (1-427)	
QY	47 ATGGGACCCCAAGCAGCGNCTCCCTGGCTGCTGCGAGCTGGACCTGGGGCAA 106
Db	1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGln 20
QY	107 CTGGAGGGCGTGGCTGGGTGAACAGAGCGCAGCGCTCCGCATCCCTTGGAAAGCAC 166
Db	21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
QY	167 GGCCTACGGCAGGATGCACAGCAGGAGATTTCGGAATCTTCAGGCTGGGCGGAGGCC 226
Db	41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
QY	227 ACTGGTGATATGTTCCCGGAGGGGATAAGCCAGACCTGCCAACCCTGGAGAGGAATTTC 286
Db	61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
QY	287 CGCTCTGCCCTCAACCGCAAGAGGGTTCGTTTAGCAGAGCAGGACCGACAGGACCTTG 346
Db	81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY	347 CACGACCCACATAAATCTACGAGTTGTGAATCTCAGGAGTTGGGACTTTTCCAGCCA 406
Db	101 HisAspProHisLysIleTyrrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY	407 GACACCTCTCGGACACCAATGGTGGAGGAGGACTTCTGATACCCAGGAAGACATCTGT 466
Db	121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
QY	467 GATGAGTTACTGGGTAACATGGTGTGGCCCCACTCCACAGATCCGGACCCCGCAAGCTG 526
Db	141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY	527 GCTGTAGCCCTGAGCCCTGCGCTCAGCCCTGCGGAGCCCGACCTTGGACAATCCCACT 586
Db	161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY	587 CCCTTCCCAACCTCGGGGCCCTCTGAGAACCACTGAAGCGGCTGTGGTCCCGGGGAA 646
Db	181 ProPheProAsnLeuGlyProSerGlnAsnProLeuLysArgLeuLeuValProGlyGlu 200
QY	647 GAGTGGGAGTTCCGAGGTGCACAGCTTCTACCGGGCGCGCCCAAGTCTTCCACAGACCATC 706
Db	201 GluTrpGluPheGluValThrAlaPheTyrrArgGlyArgGlnValPheGlnGlnThrIle 220
QY	707 TCCTGCGCGGAGGGCTCGGCTGGTGGGTCGAGTGGGAGGAGGAGGAGGAGGAGGAGGAG 766
Db	221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY	767 TGGCCAGTCACACTCCACAGACCTGGCATGTCCCTGCACAGACAGGGGAGTGCATGAGCTAC 826
Db	241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyrr 260
QY	827 GTGAGGCATGTGCTGAGCTGCTGGTGGGAGTGGCTCTCTGCGGCGCGGGCAGTGG 886
Db	261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY	887 CTCTGGGCCCGCGGCTGGGGGACTGCCACATACTGGGCACTGAGGAGGAGTGTCTC 946

Db	281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrrAlaValSerGluGlnLeuLeu 300
QY	947 CCCAACACGGCGCATGGCCCTGATGGCGAGGTCCCAAGGCAAGCAAGGAGGCGCTGTTT 1006
Db	301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
QY	1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTTCACGGAAGGAGCGGACCTCACA 1066
Db	321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
QY	1067 CGCTATGCCCTCTGTTCTGTGTGGGAGTGTATGCCCGCAGCAGCAGCGCTGGACCAAG 1126
Db	341 ArgTyrrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
QY	1127 AGGCTCGTGTGTCAGAGTTGTCGCCAGCTCCCTCAGGCGCTTGGTAGAAATGGCCCGG 1186
Db	361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
QY	1187 GTAGGGGGTCCCTCCCTCGGAGAACTGTGGACCTGCACATTTCCAAACAGCCACCCA 1246
Db	381 ValGlyGlyValAsnSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
QY	1247 CTCTCCCTCACCTCCGACCCAGTACAGGCTTACCTGCAGGACTTGTGGGCGCATGGAT 1306
Db	401 LeuSerLeuThrSerAspGlnTyrrLysAlaTyrrLeuGlnAspLeuValGluGlyMetAsp 420
QY	1307 TTCAGGGCGCTGGGAGGAC 1327
Db	421 PheGlnGlyProGlyGluSer 427
RESULT 2	
US-09-417-540-16	
; Sequence 16, Application US/09417540	
; Patent No. 6639052	
; GENERAL INFORMATION:	
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,	
; Jian Ni and Jing-Shan Hu	
; TITLE OF INVENTION: Human Genes, Sequences and	
; Expression Products	
; NUMBER OF SEQUENCES: 22	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,	
; CECCHI, STEWART & OLSTEIN	
; STREET: 6 BECKER FARM ROAD	
; CITY: ROSELAND	
; STATE: NEW JERSEY	
; COUNTRY: USA	
; ZIP: 07068	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: 3.5 INCH DISKETTE	
; COMPUTER: IBM PS/2	
; OPERATING SYSTEM: MS-DOS	
; SOFTWARE: WORD PERFECT 5.1	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/09/417,540	
; FILING DATE: 14-Oct-1999	
; CLASSIFICATION: <Unknown>	
; PRIOR APPLICATION DATA:	
; APPLICATION NUMBER: 08/705,771	
; FILING DATE: August 30, 1996	
; ATTORNEY/AGENT INFORMATION:	
; NAME: MULLINS, J.G.	
; REGISTRATION NUMBER: 33,073	
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: 973-994-1700	
; TELEFAX: 973-994-1744	
; INFORMATION FOR SEQ ID NO: 16:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 427 amino acids	
; TYPE: amino acid	
; TOPOLOGY: linear	

```

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-417-540-16

```

Alignment Scores:

Pred. No.:	3.82e-178	Length:	427
Score:	2326.00	Matches:	427
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	86.18%	Indels:	0
DB:	4	Gaps:	0

US-09-975-253-1 (1-1426) x US-09-417-540-16 (1-427)

47	QY	ATGGGAACCCAAAGCCACGGNCTCTGCCCTGGCTGTGTGTCGACGTGCACCTGGGGCAA	106
1	Db	MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln	20
107	QY	CTGAGGCGCTGGCTGGTGCACACAGAGCGCAGCGCTTCCGCATCCCTTGAAGCAC	166
21	Db	LeuGluGlyValAlaIrrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis	40
167	QY	GSCCTACGCGCAGGATGCACACAGAGAGATTCCGAATCTTCCAGGGCTCGCGCCGAGGCC	226
41	Db	GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaIrrpAlaGluAla	60
227	QY	ACTGCTGATATGTTCCGGAGGATGAAGCCAGACCTGCCAACCTGGAAGAGAAATTC	286
61	Db	ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe	80
287	QY	CGCTCTGCCCTCAACCCGCAAGAGGGTTGCGTTTACAGAGACCGGACCAAGCACCT	346
81	Db	ArgSerAlaLeuAsnArgLysGlnGluLeuArgLeuAlaGluAspArgSerLysAspPro	100
347	QY	CACGACCCACATAAAATCTCACAGTTTGTGAATCTCAGAGCTTCGGAGACTTTCCACGCA	406
101	Db	HisaspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro	120
407	QY	GACACTCTCCGGACACCAATGTGTGGAGGAGTACTTCTGATACCCAGAAAGACATTCG	466
121	Db	AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspIleLeu	140
467	QY	GATCAGTTACTGGGTACATGGTGTGGCCCCACTCCAGATCCGGGACCCCAAGCCTG	526
141	Db	AspGlnLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProSerLeu	160
527	QY	GCTGTAGCCCTCAGGCCCTGCCCTCAGCCCTCGGAGGCCAGCTTGACATCCCAT	586
161	Db	AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr	180
587	QY	CCCTTCCCAACTGGGGCCCTCTGAGAACCCACTGAAGCGCTGTGTGTCGGGGGAA	646
181	Db	ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu	200
647	QY	GAGTGGGAGTTCGAGGTGACAGCTTCTACCGGGCGCGAAGTCTTCAGCAGACCATC	706
201	Db	GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle	220
707	QY	TCTTCCCGGAGGCGCTCGCGCTGGTGGGTCCGAAGTCGGAGACAGAGACGCTGCTGA	766
221	Db	SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly	240
767	QY	TGGCCATCTACATGCCAGACCTGGCATGTCTCCCTGACACAGAGGGGAGTGTAGTAC	826
241	Db	TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr	260
827	QY	GTGAGGATCTGTGACCTGCTGGGTGGGGAGTGGCTCTCTGGCGGCGCGGAGTGG	886
261	Db	ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnIrrp	280
887	QY	CTCTGGGCCAGCGCTGGGGCACTGCCACACATATCGGCCAGTACGAGAGCTGCTC	946
281	Db	LeuTrpAlaGlnArgLeuGlyHisLysHisThrTyrTrpAlaValSerGlnLeuLeuLeu	300

RESULT 3

```

US-08-654-482-7
; Sequence 7, Application US/08654482
; Patent No. 6245562
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN
; TITLE OF INVENTION: MULTIPLE MYELOMA
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/654,482
; FILING DATE: 28-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 50995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-654-482-7

```

Alignment Scores:

Pred. No.:	3.71e-37	Length:	107
Score:	563.50	Matches:	105

Percent Similarity: 98.13% Conservative: 0
 Best Local Similarity: 98.13% Mismatches: 1
 Query Match: 20.88% Indels: 1
 DB: 3 Gaps: 1

US-09-975-253-1 (1-1426) x US-08-654-482-7 (1-107)

QY 65 CGGTCCTGCTGGCTGGTGTGGAGCTGGGCGCACTGGAGGGCGTGGCTGG 124
 Db 1 ArgilleLeuProTyrValSerGlnLeuAspLeuGlnLeuGluGlyValAlaTrp 20

QY 125 GTGAACAGAGCGGACGCGCTTCGCGATCCCTTGAAGCAGCGGCTACGCGAGATGCA 184
 Db 21 ValAsnLysSerArgThrArgPheArgilleProTyrLysHisGlyLeuArgGlnAspAla 40

QY 185 CAGCAGGAGGATTCGGATCTTCAGGCTGGGCGGAGCCACTGGTGCATATGTCTCC 244
 Db 41 GlnGlnGluAspPheGlyllePheGlnAlaTrpAlaGluAlaThrGlyAlaTrpValPro 60

QY 245 GGGAGGATAAGCCAGACCTGCCAACCTCGAAGAGAAATTCGCG---TCTGCGCTCAAC 301
 Db 61 GlyArgAspLysProAspLeuProThrTrpLysArgAsnPheArgSerSerAlaLeuAsn 80

QY 302 CGCAAGAGAGGTTGGTGTAGCAGAGGACCGGAGCAAGCCCTCAGCAGCCACATAAA 361
 Db 81 ArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspProHisAspProHisLys 100

QY 362 ATCTAGAGTTTGTGAACCA 382
 Db 101 IleTyrGluPheValAsnSer 107

RESULT 4
 US-09-513-999C-5034
 ; Sequence 5034, Application US/09513999C
 ; Patent No. 6783961
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Duclert, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; FILE REFERENCE: 59.US2.REG
 ; CURRENT APPLICATION NUMBER: US/09/513.999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER OF SEQ ID NOS: 3681
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 5034
 ; LENGTH: 102
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 99
 ; OTHER INFORMATION: Xaa=Asp or Glu
 US-09-513-999C-5034

Alignment Scores:
 Pred. No.: 9,2e-36 Length: 102
 Score: 546.00 Matches: 100
 Percent Similarity: 98.04% Conservative: 0
 Best Local Similarity: 98.04% Mismatches: 2
 Query Match: 20.23% Indels: 0
 DB: 4 Gaps: 0

US-09-975-253-1 (1-1426) x US-09-513-999C-5034 (1-102)

QY 47 ATGGGAACCCCAAGCAGCGNCTTCGCTGGTGTGCGAGCTGGACCTGGGGAA 106
 Db 1 MetGlyThrProLysProArgilleLeuProTyrValSerGlnLeuAspLeuGlyGln 20

QY 107 CTGGAGGCGTGGCTGGTGAACAGCCGACCGCTTCGCGATCCCTTGGAGCAC 166

Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgilleProTyrLysHis 40
 QY 167 GGCTACGGCAGGATGCACAGCAGGATTCGGAATCTCCAGGCTGGGCGGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnAspPheGlyllePheGlnAlaTrpAlaGluAla 60

QY 227 ACTGTGTCATATGTTCCCGGAGGAGTAAGCCAGACCTGCCAACCTGGAAGGAAATTC 286
 Db 61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80

QY 287 CGCTCTGCGCTCAACCGCAAGAGCGTTCGTTTTCAGCAGGACCGGAGCAAGCACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLys***Pro 100

QY 347 CACGAC 352
 Db 101 HisAsp 102

RESULT 5
 US-08-654-482-12
 ; Sequence 12, Application US/08654482
 ; Patent No. 6245562
 ; GENERAL INFORMATION:
 ; APPLICANT: Dalla-Favera, Riccardo
 ; TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN
 ; TITLE OF INVENTION: MULTIPLE MYELOMA
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/654,482
 ; FILING DATE: 28-MAY-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 50995
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 100 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-654-482-12

Alignment Scores:
 Pred. No.: 4,81e-35 Length: 100
 Score: 537.00 Matches: 98
 Percent Similarity: 98.00% Conservative: 0
 Best Local Similarity: 98.00% Mismatches: 2
 Query Match: 19.90% Indels: 0
 DB: 3 Gaps: 0

US-09-975-253-1 (1-1426) x US-08-654-482-12 (1-100)

QY 896 CAGCGCTGGGCACTGCCACATACCTGGGCGAGTGAGCGAGCTGCTCCCAACAGC 955
 Db 1 GlnArgLeuGlyHisCysHisThrTrpAlaValSerGluGluLeuProAsnSer 20

QY 956 GGGCATGGCGTGGTGGGAGTCCCGAGGACCAAGGAGGAGCGGTGTTGACCTGGGG 1015
 Db 21 GlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPheAspLeuGly 40
 QY 1016 CCTTCATTTGATGATCTTACCTTCACGGAAGGAGCGGACGCTCACACACGCTATGCC 1075
 Db 41 ProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerProArgTyrAla 60
 QY 1076 CTCTGTTTCTGTGGGGAGTCAATGGCCCGGACGAGCAGCGGTGGACAGGCTCGTG 1135
 Db 61 TrpLeuPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLysArgLeuVal 80
 QY 1136 ATGGTCAAGGTTGTGCCCGAGTGCCTCAGGCGCTTGGTGAATGCGCGGTAGGGGT 1195
 Db 81 MetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGly 100

RESULT 6

US-08-999-774A-10
 ; Sequence 10, Application US/08999774A
 ; Patent No. 6274312
 ; GENERAL INFORMATION:
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Seghezzi, Wolfgang
 ; APPLICANT: Shanahan, Frances
 ; APPLICANT: Lees, Emma M.
 ; APPLICANT: McClanahan, Terrill K.
 ; TITLE OF INVENTION: Intracellular Regulatory Molecules;
 ; TITLE OF INVENTION: Related Reagents
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DNAX Research Institute
 ; STREET: 901 California Avenue
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/999,774A
 ; FILING DATE: 10-DEC-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/032,818
 ; FILING DATE: 11-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ching, Edwin P.
 ; REGISTRATION NUMBER: 34,090
 ; REFERENCE/DOCKET NUMBER: DX0646
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650)852-9196
 ; TELEFAX: (650)496-1200
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 503 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-999-774A-10

Alignment Scores:
 Pred. No.: 1,63e-27 Length: 503
 Score: 446.00 Matches: 140
 Percent Similarity: 40.28% Conservative: 63
 Best Local Similarity: 27.78% Mismatches: 193
 Query Match: 16.52% Indels: 108
 DB: 3 Gaps: 19

US-09-975-253-1 (1-1426) x US-08-999-774A-10 (1-503)

QY 62 CCACGGNTCTGT-----CCCTGGCTGGTGTCCAGCTGGACCTGGGCAACTGGAGGGC 115
 Db 10 ProArgValLeuPheGlyGluTrpLeuLeuGlyGluIleSerSerGlyCysTyrGluGly 29
 QY 116 GTGGCTGGGTGAACAGAGCCGACGCGCTTCCGATCCCTTGGAGACACGCGCTACGG 175
 Db 30 LeuGlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpLysHisPheAlaArg 49
 QY 176 CAGGATGCACAGCAGGAGGATTCCGAATCTTCCAGGCTGGCCGAGGACCTGGTGCA 235
 Db 50 LysAspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArg 69
 QY 236 TATGTTCC-----GGAGGGATAAGCCAGACCTGCCAAC----- 271
 Db 70 TrpProProSerSerArgGlyGlyProProGluAlaGluThrAlaGluArgAla 89
 QY 272 ---TGAAGAGGAATTCGCTCTGCCCTCAACCGCAAGAGGGTTCGTTTACGAGAG 328
 Db 90 GlyTrpLysThrAsnPheArgCysAlaLeuArgSerThrArgArgPheValMetLeuArg 109
 QY 329 GACCGGACGACGACCCCTCACGACCCACATAAAATCTACAGCTTCTG----- 376
 Db 110 AspAsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCys 129
 QY 377 ---AACTCAGGAGTTGG-----GACTTTTCCAGCCAGACACCTCTCCGGAC 421
 Db 130 TrpArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaValProPro 149
 QY 422 ACCAATGGTGA----- 433
 Db 150 ProGlnGlyGlyProProArgProPheLeuAlaHisThrProAlaGlyLeuGlnAlaPro 169
 QY 434 -----GGCAGTACTTCTGAT-----ACCCAG 454
 Db 170 GlyProLeuProAlaProAlaGlyAspLysGlyLeuLeuGlnAlaValGlnGln 189
 QY 455 GAAGACATTTGGATGAGTTACTG---GGTAACATGGTGTGGCCCTCCCTCCCA----- 505
 Db 190 SerCysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLys 209
 QY 506 -----GATCCGGGA 514
 Db 210 AlaProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGly 229
 QY 515 CCCCCA-----AGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGCGCTCCCTCAG 553
 Db 230 LeuProAlaGlyGluLeuTyrGlyTrpAlaValGluLysThrProSerProGlyProGln 249
 QY 554 CCGCTGGGAGCCCCAGCTTGGACAATCCCACTCC-----TTCCCA 595
 Db 250 ProAlaAlaLeuThrGlyGluAlaAlaProGluSerProHisGlnAlaGluPro 269
 QY 596 AACCTGGGGCCCTCTGAGAACCCACTGAAGCGCTGTGGTGGCGGGGGAAGTGGAG 655
 Db 270 TyrLeuSerProSerProSerAlaCysThrAlaValGlnGluPro---SerProGlyAla 288
 QY 656 TTCAGGTGACAGCTTCTACCGGGCGCCCACTCTCCAGCAGACCATCTCTCTGCCCC 715
 Db 289 LeuAspValThrIleMetTyrLysGlyArgThrValLeuGlnLysValValGlyHisPro 308
 QY 716 GAGGCGCTGGCGTGGTGGG---TCCGAAGTGGGAGACAGGACGCTGCTGGATGGCA 772
 Db 309 SerCysThrPheLeuTyrGlyProProAspProAlaValArgAlaThrAspProGlnGln 328
 QY 773 GTACACTGCCAGACCTGGCATGCTCCCTCAGACAGGGGAGTGTAGTACGTAGG 832
 Db 329 ValAlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrThrGlu 347
 QY 833 CATGTCTGAGTGCCTGGGTGGGGGACTGCTCTCTGGCGGGCGGCGAGTGGCTCTGG 892
 Db 348 GluLeuLeuArgHisValAlaProGlyLeuHisLeuGluLeuArgGlyProGlnLeuTrp 367

Score:	397.00	Matches:	125
Percent Similarity:	41.63%	Conservative:	54
Best Local Similarity:	29.07%	Mismatches:	187
Query Match:	14.71%	Indels:	64
DB:	2	Gaps:	16
US-09-975-253-1 (1-1426) x US-08-611-280-25 (1-450)			
QY	77	TGGTGGTGTGCGACGTGGACCTGGGCAACTGGAGCGCTGGCTGGTGAACAGAGC	136
DB	27	TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGluGlu	46
QY	137	CGCAGCGCTTCCGATCCCTTGGAGCAGCGCCTACGCCAGGAT---GCACAGCAGGAG	193
DB	47	LysSerIlePheArgIleProTrpIysHisAlaGlyLysGlnAspTyrAsnArgGluGlu	66
QY	194	GATTTCGGAATCTTCAGCGCTGGCGGCGGACCTGGTGCATATGTTCCCGGGAGGAT	253
DB	67	AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeuAsp	86
QY	254	AAGCCAGACCTGCCCAACCTGGAGAGGAATTTCCGCTCTGCCCTCAACCGCAAGAGG	313
DB	87	LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp	106
QY	314	TTGCGTTTAGCAGAGACCGGAGCAAG---GACCCCTCAGCAGCCACATAAATCTACGAG	370
DB	107	PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg	126
QY	371	TTTGTGAACCTAGGAGTTGGGACATTTCCAGCCAGACACACCTCTCCGACACCAATGGT	430
DB	127	IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuLysAspProGlnMet	146
QY	431	GGAGGCGAGT---ACTTCTGATACCCAGGAGACATTTCTGGATGAGTTACTGGGT	481
DB	147	SerMetSerHisProTyrThrMetThrThrProTyrProSerLeuProAlaGlnValHis	166
QY	482	AACATGGTGTGGCCCGCCACTC---CCAGATCCCGGAGGCC	517
DB	167	AsnTyrMetMetProProLeuAspArgSerTrpArgAspTyrValProAspGlnProHis	186
QY	518	CGAGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGCCCTCAGCCCTG---559	
DB	187	ProGluIle---ProTyrGlnCysProMetThrPheGlyProArgGlyHisHis	203
QY	560	---CGGAGCCCGCAGTTGGACAAAT---CCCACCTCCCTTCCCAACCTGGG	604
DB	204	TrpGlnGlyProAlaCysGluAsnGlyCysGlnValThrGlyThrPheTyrAlaCysAla	223
QY	605	CCCTCTGAGAAC---CCACTGAAGCGGCTGTGGTGGCGGGAAGAG	649
DB	224	ProProGluSerGlnAlaProGlyValProThrGluProSerIleArgSerAlaGluAla	243
QY	650	TGGAGTTC---GAGGTGACAGCCTTCTACCGGGGCCCGCAAGTCTTC	694
DB	244	LeuAlaPheSerAspCysArgLeuHisIleCysLeuTyrTyrArgGluIleLeuValLys	263
QY	695	CAGCAGACCATCTCTCCCGGAGGCGCTCGCGCTGGTGGGTCGCGAGTGGGAGACAGG	754
DB	264	GluLeuThrThrSerSerProGluGlyCysArgIleSerHisGlyHisThrTyrAspAla	283
QY	755	ACGTGCTGATGGCCAGTCACACTGCCAGACCCCTGGCATGTCCCTGACACAGGCGGA	814
DB	284	SerAsnLeuAsp---GlnValLeuPheProTyrPro---GluAspAsnGly	298
QY	815	GTGATGAGCTACGTGAGGAGCATGTCTGAGCTGCGCTGGTGGGAGGACTGGCTCTCTGGCGG	874
DB	299	GlnArgLysAsnIleGluLysLeuLeuSerHisLeuGluArgGlyValValLeuTrpMet	318
QY	875	GCCGGGCGAGTGGCTCTGGGCGCCAGCGGTGGGGGACTCTCCACACACTACTGGGCATGAGC	934
DB	319	AlaProAspGlyLeuTyrAlaLysArgLeuCysGlnSerArgIleTyrTrp---335	
QY	935	GAGGAGCTGTCTCCCAACAGCGGGCATGGGCT---GATGGC	973

QY	893	GCCAGCGCTGGGCGACTGTCACACATATCTGGGCGAGTGGAGGAGCTCTCCCAAC	952
DB	368	AlaArgArgMetGlyLysCysLysValTyrTrpGluValGly-----381	
QY	953	AGCGGCGATGGCGCTGATGGCGAG-----GTCCCAAG	985
DB	382	-----GlyProGlySerAlaSerProSerThrProAlaCysLeuLeuProArg	398
QY	986	GACAAGAGGAGGCGGTGTTGACCTGGGCGCCTTCATTGTAGATCTGATTACCTTCAGC	1045
DB	399	AsnCysAspThrProIlePheAspPheArgValPhePheGlnGluLeuValGluPheArg	418
QY	1046	GAAGAAAGCGGAGCG---TCACAGCCTATGCCCTCTGTTCTGTGGGAGTCAATCG	1102
DB	419	AlaArgGlnArgGlySerProArgTyrThrIleTyrLeuGlyPheGlyGlnAspLeu	438
QY	1103	CCCCAGGACCGCTGGACAGGCTGCTGATGGTCAAGTGTGTGCCACGTCGCTC	1162
DB	439	SerAlaGlyArgProLysGlyLysSerLeuValLeuValLysLeuGluProTrpLeuCys	458
QY	1163	AGGGCGCTTGTAGAAATGCCCGGTAGGGGTGCTCTCCCTGGAG---AATACTGTG	1219
DB	459	ArgValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerSerLeu	478
QY	1220	GACCTGCATTTCCAAAGCCACCCACTCTCCCTCACCTCCGACAGTACAGGCTAC	1279
DB	479	SerLeuCysLeuSerSerAlaAsnSerLeu-----TyrAspAspIleGluCysPhe	495
QY	1280	CTGAGGACTTG 1291	
DB	496	LeuMetGluLeu 499	
RESULT 7			
US-08-611-280-25			
Sequence 25, Application US/08611280			
Patent No. 5891666			
GENERAL INFORMATION:			
APPLICANT: Matsuyama, Toshifumi			
APPLICANT: Grossman, Alex			
APPLICANT: Richardson, Christopher D.			
TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES			
NUMBER OF SEQUENCES: 25			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Angen Canada Inc.			
STREET: 6733 Mississauga Road, Suite 303			
CITY: Mississauga			
STATE: Ontario			
COUNTRY: Canada			
ZIP: L5N 6J8			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent In Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/611,280			
FILING DATE:			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: Oleski, Nancy A.			
REGISTRATION NUMBER: 34,688			
REFERENCE/DOCKET NUMBER: A-338A			
INFORMATION FOR SEQ ID NO: 25:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 450 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: protein			
US-08-611-280-25			
Alignment Scores:	1.32e-23	Length:	450
Pred. No.:			

336 -----AspGlyProLeuAlaLeuCysAsnAspArgProAsn 347
 QY 974 GAGTCCCAAGACAGGAGGCGTGTGACCTGGGGCCCTTCATTTGATCTG 1033
 Db LysLeuGluArgAspGlnThrCysLysLeuPheAspThrGlnGlnPheLeuSerGluLeu 367
 QY 1034 ATTACCTTCACGGAAGACGACGCTACACAGCTATGCCCTCTGGTTCGTGTGGG 1093
 Db GlnAlaPheAlaHisGlyArgSerLeuProArgPheGlnValThrLeuCysPheGly 387
 QY 1094 GAGTCATGGCCCGACGACGACGCTGGACCAAGAGGCTCGTATGTCAGGTTGTGCC 1153
 Db GluGluPheProAspProGlnArg---GlnArgLysLeuIleThrAlaHisValGluPro 406
 QY 1154 ACCTGCTCAGGCGCTGTAGAAATGGCCCGGTAGGGGTGCCCTCCCTGGAGAAT 1213
 Db LeuLeuAlaArgGlnLeuTyrrPheAlaGlnGlnAsnSerGlyHisPheLeuArgGly 426
 QY 1214 ACTGTGGACTG-----CACATTTCCAC 1237
 Db -----TyrAspLeuProGluHisIleSerAsn 435

RESULT 8

US-09-195-940-25
 ; Sequence 25, Application US/09195940
 ; Patent No. 6258935
 ; GENERAL INFORMATION:
 ; APPLICANT: Matsuyama, Toshifumi
 ; APPLICANT: Grossman, Alex
 ; APPLICANT: Richardson, Christopher D.
 ; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amgen Canada Inc.
 ; STREET: 6733 Mississauga Road, Suite 303
 ; CITY: Mississauga
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: L5N 6J8
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: US/09/195,940
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/611,280
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oleski, Nancy A.
 ; REGISTRATION NUMBER: 34,688
 ; REFERENCE/DOCKET NUMBER: A-338A
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 450 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-195-940-25

Alignment Scores:
 Pred. No.: 1.32e-23 Length: 450
 Score: 397.00 Matches: 125
 Percent Similarity: 41.63% Conservative: 54
 Best Local Similarity: 29.07% Mismatches: 187
 Query Match: 14.71% Indels: 64
 DB: 3 Gaps: 16

US-09-975-253-1 (1-1426) x US-09-195-940-25 (1-450)
 QY 77 TGCTGTGTGTCAGCTGGACCTGGGGCAACTGGAGGGCTGGCTGGGTGAACAAGAGC 136
 Db TTPLeuIleAspGlnIleAspSerGlyLysTyrrProGlyLeuValTrpGluAsnGluGlu 46
 QY 137 CGACACGGCTTCGCTATCCCTTGGAAAGCAGCGGCTACGGCAGAT---GCACAGCAGGAG 193
 Db LysSerIlePheArgIleProTrpLysHisAlaGlyLysGlnAspTyrrAsnArgGluGlu 66
 QY 194 GATTTCGGATCTTCAGGCTGGGGCGGAGGCACTGTCATATGTTCCCGGAGGGAT 253
 Db AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyIleAsp 86
 QY 254 AAGCCAGACCTGCCAACCTTGAAGAGGAATTTCCCTCTGCCCTCAACGCAAGAGGG 313
 Db LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
 QY 314 TTGCGTTTACGAGAGGACCGGAGCAAG---GACCCCTACAGCCACATAAATCTACGAG 370
 Db PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTrpLysValTyrrArg 126
 QY 371 TTTGTGAACCTCAGGAGTTGGGGACTTTTCCAGGCAGACACCTCTCCGACACCAATGTT 430
 Db IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuGluAspProGlnMet 146
 QY 431 GGAGGCAGT-----ACTTCTGATACCCAGGAAGACATTTCTGGATCAGTTACTGGGT 481
 Db SerMetSerHisProTyrrThrMetThrProTyrrProSerLeuProAlaGlnValHis 166
 QY 482 AACATGGTGTGGCCCTCCTC-----CCAGATCCGGGAGCC 517
 Db AsnTyrrMetMetProLeuAspArgSerTrpArgAspTyrrValProAspGlnProHis 186
 QY 518 CCAAGCTGGGTAGCCCTCCTCAGCCCTCCTCAGCCCTC-----559
 Db ProGluIle-----ProTyrrGlnCysProMetThrPheGlyProArgGlyHisHis 203
 QY 560 ---CGAGCCCTCTGGACAAT-----CCACTCTCTCCCAACCTGGGG 604
 Db TrpGlnGlyProAlaCysGluAsnGlyCysGlnValThrGlyThrPheTyrrAlaCysAla 223
 QY 605 CCCTCTGAGAAC-----CACTGAAGCGCTGTGGTCCCGGGGGAAGAG 649
 Db ProProGluSerGlnAlaProGlyValProThrGluProSerIleArgSerAlaGluAla 243
 QY 650 TGGGAGTTC-----GAGGTGACAGCTTCTACCGGGGCGCCCAAGTCTTC 694
 Db LeuAlaPheSerAspCysArgLeuHisIleCysLeuTyrrArgGluIleLeuValLys 263
 QY 695 CAGCAGACCATCTCTGCTCCCGAGGGCTGGGTGGTGGGTCCGAAAGTGGGACAGG 754
 Db GluLeuThrThrSerSerProGluGlyCysArgIleSerHisGlyHisThrTyrrAspAla 283
 QY 755 ACGTGCCTGATGCCAGTCACATGCCAGACCTGGCATGCTCCCTGACAGACAGGGA 814
 Db SerAsnLeuAsp---GlnValLeuPheProTyrrPro-----GluAspAsnGly 298
 QY 815 GTGATGAGCTACGTGAGGCATGCTGAGCTGCTGGGTGGGTGGGTGGGTCTCTGGCGG 874
 Db GlnArgLysAsnIleGluLysLeuLeuSerHisLeuGluArgGlyValValLeuTrpMet 318
 QY 875 GCCGGCAGTGGCTTGGGGCCAGCGCTGGGGCAGCTGCCACATACTGGCGCAGTGAGC 934
 Db AlaProAspGlyLeuTyrrAlaLysArgLeuCysGlnSerArgIleTyrrTrp-----335
 QY 935 GAGGAGCTGCTCCCAACAGCGGGCATGGGCT-----GATGGC 973
 Db -----AspGlyProLeuAlaLeuCysAsnAspArgProAsn 347
 QY 974 GAGTCCCCAAGACAGGAGGCGGTGTTTACCTGGGGCCCTTCATTTGATCTG 1033
 Db LysLeuGluArgAspGlnThrCysLysLeuPheAspThrGlnGlnPheLeuSerGluLeu 367

1034 ATTACCTTCACGAGGAGCGGACGCTCACACGCTATGCCCCTGCTGTTCTGTGTGGG 1093
1094 GAGTCATGCCCCAGGACGAGCGGTCGACGAGGCTCGTATGCTCAAGGTTGTGCCC 1153
1154 ACGTGCCTCAGGCGCTCTGAGAAATGCGCGGTAGGGGTGCTCCTCCTCGGAGAA 1213
1214 ACTGTGGACCTG-----CACATTCCCAAC 1237
427 ---TyrAspLeuProGluHisSerAen 435

US-09-562-466-25
Sequence 25, Application US/09562466
Patent No. 6369202
GENERAL INFORMATION:
APPLICANT: Matsuyama, Toshifumi
Richardson, Christopher D.
TITLE OF INVENTION: NOVEL GENES ENCODING LSTRF POLYPEPTIDES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSER: Amgen Canada Inc.
STREET: 6733 Mississauga Road, Suite 303
CITY: Mississauga
STATE: Ontario
COUNTRY: Canada
ZIP: L5N 6J8
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/562,466
FILING DATE: 01-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/195,940
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-562-466-25

Alignment Scores:
Pred. No.: 1.32e-23 Length: 450
Score: 397.00 Matches: 125
Percent Similarity: 41.63% Conservative: 54
Best Local Similarity: 29.07% Mismatches: 187
Query Match: 14.71% Indels: 64
Gaps: 16

US-09-975-253-1 (1-1426) x US-09-562-466-25 (1-450)
77 TGGCTGGTGTCCAGCTGACCTGGGCAACTGGAGGCGTGGCTGGGTGACAAAGAGC 136
27 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGlu 46

QY 137 CGACGCGCTTCGCTATCCCTTGAAGCAGCGGCTACGGCAGGAT---GCACAGCAGAG 193
Db 47 LysSerIlePheAerGlieProTriPheHisAlaGlyLysGlnAspTyrAsnArgGluGlu 66
QY 194 GATTTCGGAATCTTCAGGCTGCGCGGAGCCACTGTCATATGTTCCCGGAGGAT 253
Db 67 AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLys 86
QY 254 AAGCCAGACCTGCGCAACCTGGAAGGAAATTTCCGCTCTGCGCTCAACCGCAAGAGGG 313
Db 87 LysProAspPropToTrpThrPheLysThrArgLeuArgCysAlaLeuAsnLysSerAsn 106
QY 314 TTGCGTTTAGCAGAGGACCGGAGCAAG---GACCCCTCAGCAGCCACATATAATCTAC 370
Db 107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyr 126
QY 371 TTTGTGAACCTCAGGAGTTGGGACTTTTCCAGCCAGACACCTCTCCGACACCAATGT 430
Db 127 IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuGluAspProGln 146
QY 431 GGAGGCAGT-----ACTTCTCATACCCAGGAAGACATTTCTGGATGAGTTACTGG 481
Db 147 SerMetSerHisProTyrThrMetThrThrProTyrProSerLeuProAlaGlnVal 166
QY 482 AACATGTTGTTGCCCCCACTC-----CCAGATCGGGACCC 517
Db 167 AsnTyrMetMetProProLeuAspArgSerTrpArgAspTyrValProAspGlnPro 186
QY 518 CCAAGCTGCTGTAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGCCCTCAG 559
Db 187 ProGluIle-----ProTyrGlnCysProMetThrPheGlyProArgGlyHisHis 203
QY 560 ---CGGAGCCCGAGCTTGACAAAT-----CCCACTCCTCTCCCAACCTGGGG 604
Db 204 TrpGlnGlyProAlaCysGluAsnGlyCysGlnValThrGlyThrPheTyrAlaCys 223
QY 605 CCCTCTGAGAAC-----CCACTGAAGCGCTGTGTGTCGGGGGAGAG 649
Db 224 ProProGluSerGlnAlaProGlyValProThrGluProSerIleArgSerAlaGlu 243
QY 650 TGGGAGTTC-----GAGGTGACAGCCTTCTACCGGCGCCGCCAAGTCTTC 694
Db 244 LeuAlaPheSerAspCysArgLeuHisIleCysLeuTyrTrpArgGluLeuVal 263
QY 695 CACGACACCATCTCTGCGCGAGGCGCTGCTGGGTCGCGGTCGGAAGTGGAGACAG 754
Db 264 GluLeuThrThrSerSerProGluGlyCysArgIleSerHisGlyHisThrTyrAs 283
QY 755 ACGCTGCTGGATGGCCAGTCACACTGCCAGACCTGGCATGTCCTGCAGACAGAGG 814
Db 284 SerAsnLeuAsp---GlnValLeuPheProTyrPro-----GluAspAsnGly 298
QY 815 GTGATGAGCTACGTGAGGCATGTCTGAGCTGCTGGTGGGGAGTGGCTCTCTGCGG 874
Db 299 GlnArgLysAsnIleGluLysLeuLeuSerHisLeuGluArgGlyValLeuTrpMet 318
QY 875 GCGGGCAGTGGCTCTGCGGCCAGCGCTGGGGCACTGCCACATCACTGGCAGTACG 934
Db 319 AlaProAspGlyLeuTyrAlaLysArgLeuCysGlnSerArgIleTyrTrp----- 335
QY 935 GAGGAGTGTCTCCCAACAGCGGCGATGGGCT-----GATGGC 973
Db 336 -----AspGlyProLeuAlaLeuCysAsnAspArgProAsn 347
QY 974 GAGTCTCCCAAGGACAAAGAGGAGCGCTGTTTACCTGGGGCCCTTCATTGTAGAT 1033
Db 348 LysLeuGluArgAspGlnThrCysLysLeuPheAspThrGlnGlnPheLeuSerGlu 367
QY 1034 ATTACCTTCAGGAAGGAGCGGACGCTCACACGCTATGCGCTCTGTTCTGTGGG 1093
Db 368 GlnAlaPheAlaHisHisGlyArgSerLeuProArgPheGlnValThrLeuCysPhe 387

Fri Oct 29 11:11:44 2004

```

QY      1211 AATACCTGTGGACCTG-----CACATTTCAC 1237
Db      427 Gly---TyrAspLeuProGluHisSerAsn 436

RESULT 11
US-08-611-280-2
; Sequence 2, Application US/08611280
; Patent No. 5891666
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/611,280
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-280-2

Alignment Scores:
Pred. No.: 6.96e-23 Length: 450
Score: 388.00 Matches: 124
Percent Similarity: 40.13% Conservative: 55
Best Local Similarity: 27.80% Mismatches: 183
Query Match: 14.38% Indels: 84
DB: 2 Gaps: 17

US-09-975-253-1 (1-1426) x US-08-611-280-2 (1-450)
QY      77 TGGCTGGTGTCCAGCTGGACCTGGGCACTGGAGGCGTGGCTGGTGAACAAGAGC 136
Db      27 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGluGlu 46
QY      137 CGCAGCGCTTCGGCATCCCTTGAAGCACCGCCCTACGGCAGGAT---GCACAGCAGAG 193
Db      47 LysSerValPheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGluGlu 66
QY      194 GATTTCGGAATCTTCAGCGCTGGCCGAGCCACTGGTCATATGTTCCCGGAGGAT 253
Db      67 AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeasp 86
QY      254 AAGCCAGACCTGCCACCTGGAGAGGATTTCCGCTCTGGCTCAACCGCAAGAGAGG 313
Db      87 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
QY      314 TTGCGCTTTAGCAGACCGGAGCAAG---GACCTCTACGACCCACATAAATCTACGAG 370
Db      107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg 126

371 TTTGTG-----AACTCAGGAGTGGGACATTTTCCAGCCAGACACCTCT--- 415
Db      127 IleValProGluGlyAlaLysLysGlnLeuThrLeuAspAspThrGlnMet 146
QY      416 -----CCGGACACCAATGGTGGAGCGCAGTACTTCTGATACCCAGGAA 457
Db      147 AlaMetGlyHisProTyrProMetThrAlaProTyrGlySerLeuProAlaGlnGln--- 165
QY      458 GACATTCTGGATGAGTTACTGGTAACATGCTGTGGCCCA----- 499
Db      166 -----ValHisAsnTyrMetMetProProHisAspArgSerTrpArg 179
QY      500 -----CTCCAGATCCGGACCCCAAGCCTGGCTGTAGCCCTGAGCCCTGAGCCCTCAG 553
Db      180 AspTyrAlaProAspGlnSerHisProGluIle-----ProTyrGlnCysProVal 196
QY      554 CCCCTG-----CGAGCCCGCAGCTTGGACAAT----- 580
Db      197 ThrPheGlyProArgGlyHisTrpGlnGlyProSerCysGluAsnGlyCysGlnVal 216
QY      581 CCCACTCTCTCCCAACCTGGGGCCCTCTGAGAAC-----CCACTGAAG 625
Db      217 ThrGlyThrPheTyrAlaCysAlaProProGluSerGlnAlaProGlyIleProIleGlu 236
QY      626 CGGCTGTTGGTCCGGGGAA-----GAGTGGAGTTCGAGGTGACAGCC 670
Db      237 ProSerIleArgSerAlaGluAlaLeuAlaLeuSerAspCysArgLeuHisLeCysLeu 256
QY      671 TTCTACCGGGCGGCAAGTCTTCAGACAGACCATCTCTCCGCGAGGGCCCTGCGCTG 730
Db      257 TyrTyrArgAspIleLeuValLysGluLeuThrThrSerProGluGlyCysArgIle 276
QY      731 GTGGGTCCGAAGTGGGACAGAGCGCTGGATGGCCAGTACATGCCAGACCCCT 790
Db      277 SerHisGlyHisThrTyrAspValSerAsnLeuAsp---GlnValLeuPheProTyrPro 295
QY      791 GGCATCTCCCTGACAGACAGGGAGTGTAGTACGTAGTGGCATGTGCTGAGCTGCCTG 850
Db      296 -----AspAspAsnGlyGlnArgLysAsnIleGluLysLeuSerHisLeu 311
QY      851 GGTGGGGAGTGGCTCTCTGGCGGGCGGCGAGTGGCTGGCCAGCCAGCGCTGGGCGAC 910
Db      312 GluArgGlyLeuValLeuTrpMetAlaProAspGlyLeuTyrAlaLysArgLeuCysGln 331
QY      911 TGCCACACATCTGGCAGTGGCAGGAGGAGTCTCTCCCAACAGCGGCGATGGCCCT--- 967
Db      332 SerArgIleTyrTrp-----AspGlyProLeu 340
QY      968 -----GATGGCGAGTCCCAAGGACAAAGGAGGAGCGGTGTTGAC 1009
Db      341 AlaLeuCysSerAspArgProAsnLysLeuGluArgAspGlnThrCysLysLeuPheAsp 360
QY      1010 CTGGGGCCCTTCTATTGTAGATCTGATTACCTTCACGGAAGAGCGGACGCTCACACGC 1069
Db      361 ThrGlnGlnPheLeuSerGluLeuGlnValPheAlaHisGlyArgProAlaProArg 380
QY      1070 TATCCCTCTGTCTGTGGGGAGTCAATGGCCCAAGGACCGCGTGGACCAAGAGG 1129
Db      381 PheGlnValThrLeuCysPheGlyGluGluPheProAspProGlnArg---GlnArgLys 399
QY      1130 CTCGTGATGTGCAAGTGTGCCCCAGCTGCTCAGGGCCCTGTAGAAATGGCCCGGTA 1189
Db      400 LeuIleThrAlaHisValGluProLeuLeuAlaArgGlnLeuTyrTyrPheAlaGlnGln 419
QY      1190 GGGGTGGCTCTCCCTGGAG-----AATACCTGTGACACCTG 1225
Db      420 AsnThrGlyHisPheLeuArgGlyTyrGluLeuProGluHisValThrThrProAspTyr 439
QY      1226 CACATTTCACACAGCCAC 1243
Db      440 HisArgSerLeuArgHis 445

RESULT 12

```

; Sequence 2, Application US/09562466

Patent No. 6369202
GENERAL INFORMATION:
APPLICANT: Matsuyama, Toshifumi
Grossman, Alex
Richardson, Christopher D.
TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Canada Inc.
STREET: 6733 Mississauga Road, Suite 303
CITY: Mississauga
STATE: Ontario
COUNTRY: Canada
ZIP: L5N 6J8
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/562,466
FILING DATE: 01-May-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/195,940
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-562-466-2

Alignment Scores:
Pred. No.: 6.96e-23 Length: 450
Score: 388.00 Matches: 124
Percent Similarity: 40.13% Conservative: 55
Best Local Similarity: 27.80% Mismatches: 183
Query Match: 14.38% Indels: 84
DB: 3 Gaps: 17

US-09-975-253-1 (1-1426) x US-09-562-466-2 (1-450)

QY 77 TGGCTGGTGCAGCTGGAGCACTGGGCGCTGGCTGGTGACAAAGC 136
Db 27 TrpLeuLeuAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGluGlu 46
QY 137 CGCAGCGCTCCGATCCCTTGAAGCAGCGCTACGGCAGGAT---GCACACAGCAG 193
Db 47 LysSerValPheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGluGlu 66
QY 194 GATTTCGAATCTCCAGCTGGCGCGGCACTGGTGCATATGTTCCCGGGAGGAT 253
Db 67 AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeuAsp 86
QY 254 AAGCCAGACCTGCCAAGCTGGAAGAGGATTTCCGCTCTGCCTCAACCGCAAGAGGG 313
Db 87 LysProAspProProThrTriLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
QY 314 TTGCTTTAGCAGACCGGAGCAG---GACCTCAGCCACATATAATCTACGAG 370
Db 107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg 126
QY 371 TTTGTG-----AACTCAGGAGTTGGGACTTTTCCAGCCAGACACCTCT--- 415
Db 127 IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuAspAspThrGlnMet 146

QY 416 -----COGACACCAATGGTGGAGGAGTACTTCTGTATACCCAGAA 457
Db 147 AlaMetGlyHisProTyrProMetThrAlaProTyrGlySerLeuProAlaGlnGln--- 165
QY 458 GACATTCGTGATGAGTTACTGGGTAAACATGGTGTGGCCCA----- 499
Db 166 -----ValHisAsnTyrMetMetProProHisAspArgSerTrpArg 179
QY 500 -----CTCCAGATCCGGGACCCCAAGCTGCTGTAGCCCTGTAGCCCTGCCCTCAG 553
Db 180 AspTyrAlaProAspGlnSerHisProGluLe-----ProTyrGlnCysProVal 196
QY 554 CCCCTG-----CGAGCCCCAGCTTGGACAAT----- 580
Db 197 ThrPheGlyProArgGlyHisHisTrpGlnGlyProSerCysGluAsnGlyCysGlnVal 216
QY 581 CCCACTCCCTTCCCAACCTGGGGCCCTCTGAGAAC-----CCACTGAAG 625
Db 217 ThrGlyThrPheTyrAlaCysAlaProProGluSerGlnAlaProGlyLeuProIleGlu 236
QY 626 CGGCTGTTGGTCCGGGGGAA-----GAGTGGGAGTTCGAGGTGACAGCC 670
Db 237 ProSerIleArgSerAlaGluAlaLeuAlaLeuSerAspCysArgLeuHisIleCysLeu 256
QY 671 TTCTACCGGGCGCCAAAGTCTTCCAGCAGACCATCTCTGCCCCGAGGCGCTGGCTG 730
Db 257 TyrTyrArgAspIleLeuValLysGluLeuThrThrSerProGluGlyCysArgIle 276
QY 731 GTGGGTCCGAAGTGGGAGCAGGAGCGCTGCTGGATGCCAGTGCACACTGCCAGACCT 790
Db 277 SerHisGlyHisThrTyrAspValSerAsnLeuAsp---GlnValLeuPheProTyrPro 295
QY 791 GGCATGCTCCTGACAGACAGGCGGAGTATGAGTACGTGAGGCATGTGTGAGTCCCTG 850
Db 296 -----AspAspAsnGlyGlnArgLysAsnIleGluLysLeuLeuSerHisLeu 311
QY 851 GGTGGGGAGCTGGCTCTCTGGGGCGCGGCGAGTGGCTCTGGGCCAGCGGCTGGGGCAC 910
Db 312 GluArgGlyLeuValLeuTrpMetAlaProAspGlyLeuTyrAlaLysArgLeuCysGln 331
QY 911 TGCCACACATCTGGCGAGTGAGCGAGGAGTCTCCCAACAGCGGCGATGGCCCT--- 967
Db 332 SerArgIleTyrTrp-----GATGCGGAGTCTCCCAAGCAAGAGGAGCGCTTTGAC 1009
QY 968 -----GATGCGGAGTCTCCCAAGCAAGCAAGAGGAGCGCTTTGAC 1009
Db 341 AlaLeuCysSerAspArgProAsnLysLeuGluArgAspGlnThrCysLysLeuPheAsp 360
QY 1010 CTGGGGCCCTTCATTTAGATCTGATTACCTTCACGGAAGGAGCGGACCTCACACGC 1069
Db 361 ThrGlnGlnPheLeuSerGluLeuGlnValPheAlaHisHisGlyArgProAlaProArg 380
QY 1070 TATGCCCTCTGGTCTGTGGGGAGTCTGATGGCCAGGACCGCGTGGACCAAGAGG 1129
Db 381 PheGlnValThrLeuCysPheGlyGluGluPheProAspProGlnArg---GlnArgLys 399
QY 1130 CTCGTGATGTCAGAGTTGTGCCACCGCTGCTGAGGCGCTTGTAGAAATGGCCCGGTA 1189
Db 400 LeuIleThrAlaHisValGluProLeuLeuAlaArgGlnLeuTyrTyrPheAlaGln 419
QY 1190 GGGGGTGCCTCCTCCCTGGAG-----AATACTGTGAGCCTG 1225
Db 420 AsnThrGlyHisPheLeuArgGlyTyrGluLeuProGluHisValThrThrProAspTyr 439
QY 1226 CACATTTCCACAGCCAC 1243
Db 440 HisArgSerLeuArgHis 445

RESULT 14
US-09-230-371A-30
; Sequence 30, Application US/09230371A
; Patent No. 6348586

```

; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-30

Alignment Scores:
Pred. No.: 5,22e-20 Length: 425
Score: 352.00 Matches: 118
Percent Similarity: 42.44% Conservative: 56
Best Local Similarity: 28.78% Mismatches: 178
Query Match: 13.04% Indels: 58
DB: 3 Gaps: 14

US-09-975-253-1 (1-1426) x US-09-230-371A-30 (1-425)
QY 65 CGGNTCTCCCTGGCTGTGTCGACGCTGGGACCTGGGCAACTGGAGGGCTGGCCCTGG 124
DB 9 ArgLeuArgGlnTrpLeuGluGlnIleAspSerMetTyPrProGlyLeuIleTrp 28
QY 125 GTGAACAAGAGCGCGCTCCGCACTCCCTTGAAGCAGCGCTACGGCAGGATGCA 184
DB 29 GluAsnGluGlySerMetPheArgIleProTrpLysHisAlaGlyLysGlnAspTy 48
QY 185 CAGCAGGAG--GATTTCGAATCTTCCAGGCTGGCGGAGCCACTGGTCATATGTT 241
DB 49 AsnGlnGluValAlaSerIlePheLysAlaTrpAlaValPheLysGlyLysPheLys 68
QY 242 CCGGGAGGAGTAAGCCAGACCTGCCAACCCTGGAAGAGAAATTCGCTCTCCCTCAAC 301
DB 69 GluGly---AspLysAlaGluProAlaThrTrpLysThrArgLeuArgCysAlaLeuAsn 87
QY 302 CGAAAGAAGGTTGCGTTTAGCAGAGGACCGGAGCAAG--GACCCTCACGCCACAT 358
DB 88 LysSerProAspPheGluGluValThrAspArgSerGlnLeuAspIleSerGluProTy 107
QY 359 AAAATCTACGAGTTTGTG-----AACTCAGAGTTGGGACTTT 397
DB 108 LysValTyArgIleValProGluGluAspGlnLysCysLysLeuGlyValAlaThrAla 127
QY 398 TCCAGCCAGACACCTCCGACACCAATGGTGAGGCAGTACTTCTGAT-----ACC 451
DB 128 GlyCysValAsnGluValThrGluMetGluCysGlyArgSerGluIleAspGluLeuIle 147
QY 452 CAGGAAGACATCTGGATGATCTACTGGGTAAACATGTTGGTGGCCCACTCCAGATCCG 511
DB 148 LysGluProSerValAspAspTyMetGlyMetLleLysArgSerProSerPro----- 165
QY 512 GGACCCCAAGCTGGCTGTAGCCCTCAGCCCTGAGCCCTCAGCCCTCGGAGCCCCAGC 571
DB 166 -----ProAspAlaCysArgSerGlnLeu----- 173
QY 572 TTGGACAATCCCACTCCCTCTCCCAACCTGGGGCCCTCTGAGACCCCACTGAAGCGGCTG 631
DB 174 -----LeuProAspTrpTrpAlaHisGluProSerThrGlyArg 187
QY 632 TTGGTCCGGGGGGAAGAGTGG-----GAGTTTCAGGTG 664
DB -----

```

```

DB 188 LeuValThrGlyTyThrThrThrTyrAspAlaHisHisSerAlaPheSerGlnMetValIle 207
QY 665 ACAGCCTTCTACCGGGCGCCCAAGTCTTCCACACACATCTCTCCCGGAGGCGCTG 724
DB 208 SerPheTyThrGlyLysLeuValGlyGlnAlaThrThrThrCysProGluGlyCys 227
QY 725 CGGCTGGTGGGTCCGAAGTGGGAGACAGGACGCTCCCTGGATGGCCAGTCACTGCCA 784
DB 228 ArgLeuSerLeuSerGlnProGly-----LeuProGlyThrLysLeuTyGlyPro 244
QY 785 GACCCCTGGCATGTCCTG-----ACACAGCAGGAGGAGTGTAGCTACGCTG 829
DB 245 Glu---GlyLeuGluLeuValArgPheProProAlaAspThrIleProSerGluArgGln 263
QY 830 AGGCATGTGCTGAGCTGCTGGTGGGACTG-----GCTCTCTGGCGGCGCGG 880
DB 264 ArgGlnValThrArgLysLeuPheGlyHisLeuGluArgGlyValLeuHisSerSerArg 283
QY 881 CAGTGGCTCTGGGCGGCGGCTGGGGCACTGCCACACATCTGGGCACTGAGCGAGGAG 940
DB 284 GlnGlyValPheValLysArgLeuCysGlnGlyArgValPheCysValValValVal 303
QY 941 CTGCTCCCCAACAGCGGCGATGGGCTGATGCGGAGTCCCAAGACACAAGAGGAGGC 1000
DB 304 ValCys-----LysGlyArgProAsnLysLeuGluArgAspGluValValGln 319
QY 1001 GTGTTTGACCTGGGCGCTTCACTGTAGATCTGATTACCTTCACGGAAGAGCGGACGC 1060
DB 320 ValPheAspThrSerGlnPhePheArgGluLeuGlnGlnPheTyAsnSerGlnGlyArg 339
QY 1061 TCACCACGCTATGCTCTGTTCTGTGTGGGGAGTCAATGGCCCCCAGGACCGCGTGG 1120
DB 340 LeuProAspGlyArgValValLeuLysCysPheGlyGluGluPheProAspMetAlaProLeu 359
QY 1121 ACCAAGAGCTCTGATGTCAGTTCAGTTCGCTGCGCCAGCTGCTCAGGCGCTTGTGTAAGATG 1180
DB 360 ArgSerLysLeuIleLeuValGlnIleGluGlnLeuTyValArgGlnLeuAlaGluGlu 379
QY 1181 GCC---CGGTAGGGGGTGGCTCTCTCTCTCT 1207
DB 380 AlaGlyLysSerCysGlyAlaGlySerVal 389

RESULT 15
US-09-230-371A-29
; Sequence 29, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-29

Alignment Scores:
Pred. No.: 5,65e-16 Length: 393
Score: 301.50 Matches: 108
Percent Similarity: 41.46% Conservative: 62
Best Local Similarity: 26.34% Mismatches: 177
Query Match: 11.17% Indels: 63

```

DB: 3 Gaps: 16
US-09-975-253-1 (1-1426) x US-09-230-371A-29 (1-393)
Qy 38 GCGCGGACCATGGGAACCCCAAGACCGGTCCTCCCTGGCTGGTGTGCAGCTGGAC 97
Dy 4 GlyArgAlaArgCysThrArgLysLeuArgAsn-----TTPValValGluValGlu 21
Qy 98 CTGGGCAACTGGAGGGCGTGGCTGGTGAACAGCCGACGCGCTTCCGATCCCT 157
Dy 22 SerGlyGlnPheProGlyValCysTrpAspThrAlaLysThrMetPheArgIlePro 41
Qy 158 TGAAGACACGGCTACCGCAGAT---GCACAGCAGGAGGATTTCCGAACTTCCAGGCC 214
Dy 42 TrpLysHisAlaGlyLysGlnAspPheArgGluAspGlnAspAlaAlaPhePheLysAla 61
Qy 215 TGGCGCGAGGCCACTGGTGATATCTCCCGGAGGGATAAGCCAGACCTGCCAACCTGG 274
Dy 62 TrpAlaIlePheLysGlyLysTrpLysGluGly---AspThrGlyGlyProAlaValTrp 80
Qy 275 AAGAGGAATTTCCGCTCTGCCCTCAACCGCAAGAGGGTTCCGTTAGCAGAGGACCGG 334
Dy 81 LysThrArgLeuArgCysAlaLeuAsnLysSerSerGluPheLysGluValProGluArg 100
Qy 335 AGCAAG---GACCCTCACGCCACCAATAAATCTACGAGTTGTGNACTCAGAGTTGGG 391
Dy 101 GlyArgMetAspValAlaGluProTyrLysValTyrGlnLeuLeuProProGlyIleVal 120
Qy 392 GACTTTTCCAGCCAGACACTCT-----CCGACACCAATGGTGGAGGAGTACTTCT 445
Dy 121 Ser---GlyGlnProGlyThrGlnLysValProSerLysArgGlnHisSerSerValSer 139
Qy 446 GATACCCAGCAGACATCTGGATGAGTTACTGGGTAACATGTGTGGTGGCCCACTC--- 502
Dy 140 SerGluArgLysGlu---GluGluAspAlaMetGlnAsnCysThrLeuSerProSerVal 158
Qy 503 -----CCAGAT 508
Dy 159 LeuGlnAspSerLeuAsnGluGluGluGlyAlaSerGlyGlyAlaValHisSerAsp 178
Qy 509 CCGGGACCCCAAGCTGGCTGTAGCCCTGAGCCCTGCGCTCCCTCAGCCCTTGGCGGCC 568
Dy 179 IleGlySerSerSerSerSerProGluPro-----Gln 191
Qy 569 AGCTTGACAACTCCACTCCCTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGACGG 628
Dy 192 GluValThrAspThrThrGluAlaProPheGlnGly---AspGlnArgSerLeuGluPhe 210
Qy 629 CTGTTGGTGGCGGGAAGAGTGGGAGTTGAGGTGACAGCTTCTACCGGGCGCCCAA 688
Dy 211 LeuLeuProGluProAspTyrSerLeuLeuLeuThrPheIleTyrAsnGlyArgVal 230
Qy 689 GTCTTCAG-----CAGACCATCTCTCCCGGAGGCGCTGCGGCTGGTGGGTCC 739
Dy 231 ValGlyGluAlaGlnValGlnSerLeuAspCys-----ArgLeuValAlaGlu 246
Qy 740 GAAGTGGGAGACAGACGCTGCTGGATGGCCAGTCACATGCGCAGACCTGGCATGTCC 799
Dy 247 ProSerGlySerGluSerSerMetGlu---GlnValLeuPheProLysPro----- 262
Qy 800 CTGACACAGGAGGAGTGTAGTACGTGAGGCTGTGCTGAGCTGCTGGTGGGGA 859
Dy 263 -----GlyProLeuGluProThrGlnArgLeuLeuSerGlnLeuGluArgGly 278
Qy 860 CTGGCTCTCTGGCGGCGGCGGAGTGGCTCTGGGCCAGCGCTGGGCACTGCCACACA 919
Dy 279 IleLeuValAlaSerAsnProArgGlyLeuPheValGlnArgLeuCysProIleProIle 298
Qy 920 TACTGGCGAGTGGAGGAGCTGTCTCCCAACCGCGGCGATGGCCCTGATGGCGAG--- 976
Dy 299 SerTrp-----AsnAlaProGlnAlaProProGlyProGly 310
Qy 977 -----GTCCCAAGCAAGGAGGAGCGGTGTTGACCTGGGGCCCTTCATTGTA 1027

311 ProHisLeuLeuProSerAsnGluCysValGluLeuPheArgThrAlaTyrPheCysArg 330
1028 GATCTGATTACCTTCACGGAAGGAGCGGCTCAGCAGCTATGCCCTCTGTTCTGT 1087
331 AspLeuValArgTyrPheGlnGlyLeuGlyProProLysPheGlnValThrLeuAsn 350
1088 GTGGGGGAGTTCATGGCCCCCAGGACAGCCGCTGGACCAAGAGCTCGTATCGTCAAGTT 1147
351 PheTrpGluGluSerHisGlySerSerHisThrProGlnAsnLeuIleThrValLysMet 370
1148 GTGCCACAGTCCCTCAGGCGCTTGGTAGAA 1177
371 GluGlnAlaPheAlaArgTyrLeuLeuGlu 380

Search completed: October 28, 2004, 15:44:18
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 15:15:00 ; Search time 149 Seconds
(without alignments)
6205.795 Million cell updates/sec

Title: US-09-975-253-1
Perfect score: 2699
Sequence: 1 ggttcagctgcccgaagc.....aaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 2741442

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xl
-Q=/cgn2_1/USPTO spoop/US09975253/runat 28102004 111511 19667/app query.fasta_1.1607
-DB=Published Applications AA -QMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09975253 @CGN 1 1 224 @runat 28102004 111511 19667
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	2326	86.2	427	9	US-09-975-253-2	Sequence 2, Appli
2	2326	86.2	427	14	US-10-453-478-16	Sequence 16, Appli
3	2018	74.8	392	15	US-10-276-774-2333	Sequence 2333, Ap
4	399.5	14.8	94	15	US-10-258-666-31	Sequence 31, Appl
5	257	9.5	395	15	US-10-264-049-2929	Sequence 2929, Ap
c	247.5	9.2	19695	15	US-10-084-846A-3	Sequence 3, Appli
6	246	9.1	373	10	US-09-809-391-374	Sequence 374, App
7	246	9.1	373	10	US-09-882-171-374	Sequence 374, App
8	246	9.1	373	14	US-10-164-861-374	Sequence 7, Appli
9	246	9.1	373	14	US-10-084-846A-7	Sequence 46, Appl
c	243.5	8.9	19652	15	US-09-805-020-46	Sequence 8, Appli
11	240	8.9	19662	15	US-10-084-846A-6	Sequence 8, Appli
c	239.5	8.9	19662	15	US-10-084-846A-6	Sequence 356, App
c	234.5	8.8	19608	15	US-10-755-889-356	Sequence 1897, Ap
c	231.5	8.6	164	16	US-10-276-774-1897	Sequence 5, Appli
c	222.5	8.3	1008	15	US-10-084-846A-5	Sequence 141510,
c	222	8.3	19723	15	US-10-084-846A-5	Sequence 40, Appl
c	219.5	8.2	405	16	US-10-437-963-141510	Sequence 30, Appl
c	219.5	8.1	733	15	US-10-138-588-40	Sequence 40, Appl
c	218	8.1	720	14	US-10-342-331-4	Sequence 30, Appl
c	214.5	7.9	744	14	US-10-171-311-40	Sequence 30, Appl
c	214.5	7.9	744	14	US-10-236-055A-30	Sequence 37, Appl
c	214.5	7.9	744	14	US-10-301-822-37	Sequence 5, Appli
c	213	8.0	960	14	US-10-342-331-5	Sequence 184023,
c	213	8.0	19725	15	US-10-084-846A-4	Sequence 3, Appli
c	211.5	7.9	408	16	US-10-437-963-184023	Sequence 39, Appl
c	211.5	7.9	504	14	US-10-342-331-3	Sequence 10, Appl
c	211.5	7.9	744	14	US-10-156-761-11093	Sequence 56601, A
c	210.5	7.8	744	14	US-10-058-124-21	Sequence 11093, A
c	209	7.8	466	16	US-10-301-822-39	Sequence 21, Appl
c	209	7.8	1366	15	US-10-402-089-10	Sequence 1, Appli
c	209	7.8	1366	15	US-10-402-072A-10	Sequence 24, Appl
c	208.5	7.8	336	15	US-10-425-114-56601	Sequence 203630,
c	208.5	7.8	903	14	US-10-156-761-11093	Sequence 103, App
c	208.5	7.8	1078	14	US-10-058-124-21	Sequence 16, Appl
c	208	7.8	1014	14	US-10-194-441A-1	Sequence 20, Appl
c	207.5	7.8	780	15	US-10-441-965-24	Sequence 11, Appl
c	207.5	7.8	837	15	US-10-096-986-103	Sequence 6, Appli
c	207	7.7	555	16	US-10-437-963-203630	Sequence 8, Appli
c	207	7.7	1057	15	US-10-104-889-16	Sequence 10, Appl
c	207	7.7	1057	15	US-10-104-889-20	Sequence 25, Appl
c	207	7.7	1107	15	US-10-104-889-11	Sequence 261, App
c	207	7.7	1169	15	US-10-104-889-6	
c	207	7.7	1171	15	US-10-104-889-8	
c	207	7.7	1388	15	US-10-104-889-10	
c	207	7.7	1461	16	US-10-468-091-25	
c	207	7.7	1464	10	US-09-918-715-261	

ALIGNMENTS

RESULT 1
US-09-975-253-2
; Sequence 2, Application US/09975253
; Patent No. US2002016494A1
; GENERAL INFORMATION:
; APPLICANT: Paul Moore et al.
; TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
; FILE REFERENCE: PFI96P1
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 60/239,963
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)

OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Alignment Scores:
Pred. No.: 6,47e-145 Length: 427
Score: 2326.00 Matches: 427
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.18% Indels: 0
DB: 9 Gaps: 0

US-09-975-253-1 (1-1426) x US-09-975-253-2 (1-427)

QY 47 ATGGGAACCCCAAGCCAGGNTCTGCTGGCTGGTGTGCTGACCTGGGCAA 106
DB 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGln 20

QY 107 CTGGAGGGCTGGCTGGTGAACAGAGCCGCGCTTCCGCATCCCTTGGAGCAC 166
DB 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40

QY 167 GGCCTACGCGAGATGACAGCAGGAGGATTTCCGAATCTTCCAGGCTGGGCCGAGGCC 226
DB 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60

QY 227 ACTGTGTCATATGTTCCCGGGAGGGATAGCCAGACCTGCCAACCTGGNAGGAATTTC 286
DB 61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80

QY 287 CCCTCTGCCCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGACCGGAGCAAGACCTC 346
DB 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGlnAspArgSerLysAspPro 100

QY 347 CACGACCCACATAAATCTACAGATTTGTGAATCTCAGGAGTTGGGACTTTCCAGCCA 406
DB 101 HisAspProHisLysIleTrpGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120

QY 407 GACACTCTCCGACACCAATGTTGGAGGAGTACTTCTGATACCCAGGAGACATCTG 466
DB 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140

QY 467 GATGACTTACTGGTAACATGTTGGCCCACTCCAGATCCCGGACCCCAAGCCCTG 526
DB 141 AspGluLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160

QY 527 GCTGTAGCCCTCAGCCCTGCTCCAGCCCTCGGAGCCCACTTGGACAAATCCCACT 586
DB 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180

QY 587 CCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAAGCGCTGTTGGTCGGGGGA 646
DB 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200

QY 647 GAGTGGGAGTTGAGGTGACAGCTTCTACCGGGCGCCCAAGTCTTCCAGCAGACCATC 706
DB 201 GluTrpGluPheGluValThrAlaPheTyArgGlyArgGlnValPheGlnGlnThrIle 220

QY 707 TCCTGCGCGGAGGCGCTGGCGCTGGTGGGTCCGAACTGGGAGACAGGACGCTGCTGGA 766
DB 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240

QY 767 TGGCCAGTCACTGCCAGACCTTGGGATGTCCCTGACACAGAGGAGGTGATGAGCTAC 826
DB 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTrp 260

QY 827 GTGAGGATGTGCTGAGCTGCTGGGTGGGAGTCTCTGCGCGGCGCGGCGAGTGG 886
DB 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpAlaGlyGlnTrp 280

QY 887 CTCTGGGCGCGGCTGGGCACTGCCACACATCTGGGCGAGTGAGGAGGAGCTGCTC 946
DB 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTrpAlaValSerGluGluLeuLeu 300

947 CCCAACAGCGCGCATGGGCTGATGGAGGTCCCCAAGGACAAAGAGGAGGCGTGT 1006
DB 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320

QY 1007 GACCTGGGCGCTTCTGATGATCTGATGATCTGATGATCTGATGATCTGATGATCT 1066
DB 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340

QY 1067 CGCTATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
DB 341 ArgTyAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360

QY 1127 AGGCTGCTGATGTCAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186
DB 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380

QY 1187 GTAGGGGTGCT 1246
DB 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400

QY 1247 CTCTCCCTCCTCCGACGAGTACAAAGGCTTACCTGACGAGCTTGTGGAGGCGAT 1306
DB 401 LeuSerLeuThrSerAspGlnTyLysAlaTyLeuGlnAspLeuValGluGlyMetAsp 420

QY 1307 TTCCAGGCGCTGGGAGAGC 1327
DB 421 PheGlnGlyProGlyGluSer 427

RESULT 2
US-10-453-478-16
Sequence 16, Application US/10453478
Publication No. US20030208043A1
GENERAL INFORMATION:
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
Jian Ni and Jing-Shan Hu
TITLE OF INVENTION: Human Genes, Sequences and
Expression Products
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/453,478
FILING DATE: 04-Jun-2003
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705,771
FILING DATE: August 30, 1996
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 973-994-1700
TELEFAX: 973-994-1744
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-453-478-16

Alignment Scores:
 Pred. No.: 6,47e-145 Length: 427
 Score: 2326.00 Matches: 427
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.18% Indels: 0
 DB: 14 Gaps: 0

US-09-975-253-1 (1-1426) x US-10-453-478-16 (1-427)

QY 47 ATGGGAACCCAAAGCCAGGNTCTGCTCGCTGGCTGGTGTGCTGACCTGGACCTGGGCAA 106
 DB 1 MetGlyThrProLysProArg***LeuProLysLeuValSerGlnLeuAspLeuGln 20
 QY 107 CTGAGGCGCTGGCTGGTGAACAGAGCCGACGCTCCGATCCCTTTGGAAGCAC 166
 DB 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACGCGAGGATGCACAGCAGGAGATTTCGGAATCTTCCAGGCTGGGCGGAGGCC 226
 DB 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGTGCATATGTTCCCGGAGGATAGAGCCAGACCTGCCAACCTGGAAGAGGAATTTC 286
 DB 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGTCTGCCCTCAACCCGAAAGAGGTTGGCTTTAGCAGAGGACCGGAGCAAGACCCCT 346
 DB 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATATAATCTACGAGTTTGTGAATCAGAGTTGGGACCTTTTCCAGGCA 406
 DB 101 HisaspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCGGACCAACATGTTGAGCGAGTACTTCTGTATACCCAGGAGACATTTCTG 466
 DB 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspLeu 140
 QY 467 GATGAGTTACTGGGTAACATGTTGTGCCCCACCTCCAGATCCGGGACCCCCAGCCCTG 526
 DB 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCTCAGCCCTGCGGAGCCCGGAGTGGACATCCCACT 586
 DB 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGGAAGCGGCTGTTGGTCCGGGGAA 646
 DB 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTGAGGTGACAGCCTTCTACCGGGGCCGCCAAGTCTTCCAGCAGACCATC 706
 DB 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCGCGAGGCGCTGCGCTGCTGGGTCGGAAGTGGGAGCAGCAGCCTGCTCGA 766
 DB 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGACAGACCTGGCATGTCCCTGACAGACAGGGGAGTGTATGAGTAC 826
 DB 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGCAATGCTGAGTGGCTGGGTGGGAGCTGGCTCTCTGCGGGCGCGGCGAGTGG 886
 DB 261 ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCGGCGCTGGGCGACTGCCACATCTGGCAGTGTGAGGAGCAGCTGCTC 946
 DB 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluLeuLeu 300
 QY 947 CCCAACGCGGCGATGGGCGCTGATGGCGAGGTCCCAAGGACCAAGGAAGGAGGCGTGT 1006

DB 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGlyGlyValPhe 320
 QY 1007 GACCTGGGCGCTTCAATTGTAGATCTGATTACCTTACCGAAGGAGCGGACGCTCACA 1066
 DB 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGCTGTTCTGTGGGGAGTCAATGGCCCCAGGACCCGCTGGACCAAG 1126
 DB 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTCGTGATGGTCAAGTTGTGCCACGCGCTCAGGGCTTGGTGAAGATGCCCGG 1186
 DB 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGGCTCTCTCCCTGGAGATACTGTGGACCTGCACATTTCCAAACAGCCACCA 1246
 DB 381 ValGlyGlyAlaSerSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCTCGGACCATACAGGCCTACCTGAGGACTTGGTGGAGGCGATGGAT 1306
 DB 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGCGCTCGGGGAGAGC 1327
 DB 421 PheGlnGlyProGlyGluSer 427

RESULT 3
 US-10-276-774-2333
 ; Sequence 2333, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang, Y, Tom et al
 ; FILE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
 ; TITLE REFERENCE: 21272-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; PRIOR FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: Custom
 ; SEQ ID NO 2333
 ; LENGTH: 392
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-276-774-2333

Alignment Scores:
 Pred. No.: 1.25e-124 Length: 392
 Score: 2018.00 Matches: 372
 Percent Similarity: 98.41% Conservative: 0
 Best Local Similarity: 98.41% Mismatches: 0
 Query Match: 74.77% Indels: 6
 DB: 15 Gaps: 1

US-09-975-253-1 (1-1426) x US-10-276-774-2333 (1-392)

QY 212 GCCTGGGCGGAGGCGACTGCTGTCATATGTTCCCGGAGGAGTAAAGCCAGCTGCCAACC 271
 DB 15 AlaTrpAlaGluAlaThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThr 34
 QY 272 TGGAGAGGAGTATTTCCGCTCTGCGCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGAC 331
 DB 35 TrpLysArgAsnPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAsp 54
 QY 332 CGGAGCAGGAGCCCTCAGACCCACATAAATCTTACAGTTTGTGAATCTCAGGAGTTGGG 391
 DB 55 ArgSerLysAspProHisAspProHisLysIleTyrGluPheValAsnSerGlyValGly 74
 QY 392 GACCTTTCCGAGCAGACCTCTCCGAGCACCAATGGTGGAGGAGCTACTCTCTGATACC 451

FILE REFERENCE: 082382-000000US
CURRENT APPLICATION NUMBER: US/10/258,666
PRIOR FILING DATE: 2003-05-23
PRIOR APPLICATION NUMBER: JP 2000-126741
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: WO PCT/JP01/03700
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 31
LENGTH: 94
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: RHDH-279-1 ORF
US-10-258-666-31

Alignment Scores:
Pred. No.: 4,42e-18 Length: 94
Score: 399.50 Matches: 77
Percent Similarity: 87.23% Conservative: 5
Best Local Similarity: 81.91% Mismatches: 11
Query Match: 14.80% Indels: 1
DB: 15 Gaps: 1

US-09-975-253-1 (1-1426) x US-10-258-666-31 (1-94)

QY 1046 GAAGAAAGGAGCGCTACACCGCTATGCCCTCTGTTCTGTGGGGAGTCATGGCCC 1105
Db 2 GluGlySerArgHisSerProArgTyrThrLeuPheCysValGlySerTrpPro 21
QY 1106 CAGGACCGAGCGCTGACCAAGAGGCTCGTGTAGTGTCAAGTGTGTGCCACGTCCTCAGG 1165
Db 22 GluAspGlnProTrpValValysArgLeuValMetValValysValValProThrCysLeuLys 41
QY 1166 GCCTTGTAGTAATGCCCGGTAGGGGTGCTCTCTCTCTCGAGAAATCTGTGGACCTG 1225
Db 42 GluLeuLeuGluMetAlaArgGluGlyGlyAlaSerSerLeuLys---ThrValAspLeu 60
QY 1226 CACATTTCCAAAGCAGCCACCTCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 1285
Db 61 HisIleSerAsnSerGlnProIleSerLeuThrSerAspGlnTrpLysAlaCysLeuGln 80
QY 1286 GACTTGTGGAGGCGATGATTTCAGGGCCCTGGGGAGAGC 1327
Db 81 AspLeuValGluAspMetAspPheGlnAlaThrGlyGluThr 94

RESULT 5
US-10-264-049-2929
Sequence 2929, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133PI
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: Patent In Ver. 3.1
SEQ ID NO 2929
LENGTH: 395
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2929

Alignment Scores:
Pred. No.: 1.32e-08 Length: 395
Score: 257.00 Matches: 70
Percent Similarity: 43.17% Conservative: 28

75 AspPheSerGlnProAspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThr 94
QY 452 CAGGAAGACATCTGGATGAGTACTGGGTAACTGTTGGTGGCCCACTCCAGATCCG 511
Db 95 GlnGluAspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspPro 114
QY 512 GGACCCCAAGCTGGTGTAGCCCTCAGCCCTGAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGC 571
Db 115 GlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuLeuArgSerProSer 134
QY 572 TTGGACATCCACTCCCTTCCAAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGCTG 631
Db 135 LeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu 154
QY 632 TTGTGTCGGGGGAGAGTGGAGTTCGAGGTGACAGCCCTTCTACCGGGCGGCCCAAGTC 691
Db 155 LeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnVal 174
QY 692 TTCAGCAGACCATCTCTGCCCCGAGGCGCTGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 751
Db 175 PheGlnGlnThrIleSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp 194
QY 752 AGGAGCTGCTGGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
Db 195 ArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArg 214
QY 812 GGAGTGATGAGCTAGCTGAGGCGATGTGCTGAGCTGCTGGGTGGGGAGCTGCTCTCTGG 871
Db 215 GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrp 234
QY 872 CGGGCGGGCAGTGGCTTGGGCCCGCAGCGCTGGGCGCTGCGACATCTGACATCTGCGCAGTG 931
Db 235 ArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaVal 254
QY 932 ACAGGAGCTGCTCCCCAACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991
Db 255 SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys 274
QY 992 GAAGGAGGCGTGTTCACCTGGGGCCCTTCATTGTA-----GATCTG 1033
Db 275 GluGlyGlyValPheAspLeuGlyProPheIleValGlySerLeuGlyProAspLeu 294
QY 1034 ATTACCTCAGGAAGAGCGGACCTCAGCAGCTATGCCCTCTGTTCTGTGGGG 1093
Db 295 IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGly 314
QY 1094 GAGTCATGCCCGCAGACCGCTGACCAAGAGGCTCGTGTGTCGTCAGGTTGTGCC 1153
Db 315 GluSerTrpProGlnAspGlnProTrpThrLysArgLeuValMetValLysValValPro 334
QY 1154 ACCTGCTCAGGCGCTTGTAGAAATGGCCCGGTAGGGGTGCTCTCTCTCTGAGAAAT 1213
Db 335 ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn 354
QY 1214 ACTGTGAGCTGACATTTCCACAGCCCACTCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1273
Db 355 ThrValAspLeuHisIleSerAsnSerHisProLeuSerLeuThrSerAspGlnTrpLys 374
QY 1274 GCCTACTGAGGACTTGTGGAGGCGATGATTTCAGGGCCCTGGGGAGAGC 1327
Db 375 AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer 392

RESULT 4
US-10-258-666-31
Sequence 31, Application US/10258665
Publication No. US20040005578A1
GENERAL INFORMATION:
APPLICANT: Yamada, Yoji
APPLICANT: Sekine, Susumu
APPLICANT: Kikuchi, Yasuhiro
APPLICANT: Sakurada, Kazuhiro
APPLICANT: Kyowa Hakko Kogyo Co., Ltd.
TITLE OF INVENTION: Myocardial Cell Proliferation-Associated Genes

10508	ArgT	Val	Ser	Ser	Ala	Arg	Thr	Cys	Pro	Ser	Thr	Trp	Gly	Ser	Ser	Arg	Cys	10527
879	CGG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	CCG	823
10528	Cys	Ala	Arg	Ala	Trp	Ser	Ser	Gly	Pro	Arg	Pro	Gly	Ser	Thr	Gly	Thr		10543
822	CTC	ATC	ATC	CC	CT	CT	CT	CT	CAG	GC	ATC	GC	CAG	GC	ATC	GC	ATC	763
10544	--	Arg	Ser	Pro	Ser	Ser	Thr	Ala	Arg	Trp	Ser	Ser	Thr	Gly	Ser	Gly	Thr	10562
762	GGC	AGC	GC	TC	GT	TC	CC	CAC	TT	CGG	AAC	---	---	---	---	---	CC	733
10563	Arg	Ser	Pro	Ala	Cys	Gly	Ser	Arg	Thr	Arg	Ser	Thr	Gly	Gly	Gly	Pro	Ala	10582
732	ACC	AGC	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	697
10582	o	Pro	Ala	Ala	Ala	Gly	Arg	Ala	Arg	Pro	Gly	Ala	Gly	Ala	Pro	Gly	Arg	10600
696	TGG	AAG	ACT	TCG	CGC	GC	---	---	---	---	---	---	---	---	---	---	AG	670
10600	a	His	Arg	Pro	Gly	Gly	His	Arg	Arg	Ala	Leu	Pro	Gly	Ala	Asp	Arg	Ser	10620
669	GCT	GT	CAC	CT	CGA	ACT	CC	CAC	TC	CT	CC	CCG	---	---	---	---	GC	616
10620	g	Leu	Gly	Leu	Arg	Pro	Ser	Ala	Arg	Pro	Gln	Arg	Ala	Gly	Arg	Ala	Ala	10640
615	T	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	10660
10640	y	Val	Gly	Ala	Leu	Ala	Arg	Leu	Arg	Ser	Gly	His	Leu	Ala	Val	Pro	Arg	10660
609	GAG	GC	CC	CC	CAG	CTT	GGG	AAG	GG	AAG	---	---	---	---	---	---	---	556
10660	g	Arg	Pro	Ala	Gly	Asp	Gly	Arg	Asp	Ala	Arg	Arg	Gly	Pro	Cys	Gly	Ser	10680
555	GGC	T	AGG	CGC	AGG	CGT	CT	AG	---	---	---	---	---	---	---	---	---	526
10680	a	Gly	Arg	Gly	Arg	Ala	Val	Arg	Pro	Ala	Gly	Arg	Leu	Val	Arg	Cys	Ala	10700
525	AGC	CT	TGG	GGG	GT	CCG	GAT	CTG	GG	AGT	GGG	GC	CA	CAC	ATG	TTC	CC	466
10700	o	Gly	Pro	Gly	Thr	Arg	Asp	---	---	---	---	---	---	---	---	---	---	10717
465	AGA	AT	GT	CT	CT	CT	GGT	TAT	CAG	AAG	TAC	TG	CC	TCC	ACC	AT	TG	406
10717	o	Ala	Ala	Ala	Ala	Gly	Arg	Gln	Ala	Gly	Ala	Val	His	Ala	Trp	---	---	10733
405	GGT	TGG	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	355
10733	u	Gly	Ala	Arg	Ala	Gly	Pro	Val	Ala	Leu	Asp	Arg	Arg	Arg	Thr	Arg	---	10750
354	GGT	CT	GT	GAG	GGT	CT	TG	CT	CCG	GT	CT	CT	CT	CT	CT	CT	CT	295
10751	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	10761
294	GC	AG	CG	AG	CG	AAT	CT	CT	CT	CC	AG	TT	GG	CAG	GT	CT	GG	235
10761	y	Pro	Asn	Gly	Gln	Ala	His	Leu	Arg	Thr	Ala	Arg	Ser	Ala	Ala	Pro	Ala	10777
234	GC	ACC	AGT	GG	CT	CG	CC	CG	CC	TGG	AG	AT	TC	CG	AA	AT	CT	175
10778	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	10787
174	CGT	AG	CG	CT	GT	CT	CC	AAG	GG	AT	CGG	AG	CG	CG	GT	CT	GT	123
10787	a	Gly	Arg	Ser	Gly	Arg	Ala	Ala	Ala	Asp								

```

QY      66  CGTGGCTTTGGG-----TTCCC 49
      |||
Db      10847  uAlaGlyGluGlyArgAspMetThrGlnProProSerProThrMetCysValLeuPr 10867
QY      48  ATGG 45
      ||||
Db      10867  oTrp 10868

RESULT 7
US-09-809-391-374
; Sequence 374, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; PRIOR FILING DATE: 2001-03-16
; Current application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
; US-09-809-391-374

```

```

Alignment Scores:
Pred. No.: 6,93e-08 Length: 373
Score: 246.00 Matches: 104
Percent Similarity: 42.32% Conservative: 64
Best Local Similarity: 26.20% Mismatches: 161
Query Match: 9.11% Indels: 68
DB: 10 Gaps: 20

US-09-975-253-1 (1-1426) x US-09-809-391-374 (1-373)

QY 329 GACCGGAGCAGGAC---CCTCAGCACCCACATAAAATCTACGAGTTGTGAACCTCAGGA 385
Db 3 AspGlyThrIysGluValProMetAsnProValIysIleTyrGln----- 17
386 GTTCGGGACACTTTCACGCCAGACACCTCT-----CCGGACACCAATGTTGGAGGC 436
Db 18 ValCysAspIleProGlnProGlnGlySerIleIleAsnProGlySerThrGlySerAla 37
437 AGTACTCTCTGAT-----ACCAGAGACATCTTCGGATGAGTTACTGGGT 481
Db 38 ProTrpAspGluIysAspAsnAspValAspGluGluAspGluLeuAspGln 57
482 AACATGGTGTGGCCCACTCCAGATCCGGGACCCCAAGCCTGGTGTGACCCCTCAG 541
Db 58 SerGlnHisHisValProIleGlnAspThrPhe---ProPheLeuAsnIleAsnGlySer 76
542 CCC---TGCCCTCAGCCCTCGGAGCCCACTGGGACAATCCCACTCCC----- 589
Db 77 ProMetAlaProAlaSerValGlyAsnCysSerValGlyAsnCysSerProGluAlaVal 96
590 TTCCCAAACTGGGGCCCTCTGAG-----AACCCACTGAAGCGGTGTG 634
Db 97 TrpProLysThrGluProLeuGluMetGluValProGlnAlaProIleGlnProPheTyr 116
635 GTGCCGGGGAGACTGG-----GAGTTCGAGGTGACGCCTTC 673
Db 117 SerSerProGluLeuTrpIleSerSerLeuProMetThrAspLeuAspIleIysPheGln 136

```

```
QY 674 TACGGGCGCCCAAGTCTTCAG---CAGACCATCTCTGCGCGAGGGCTCGGGCTG 730
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 TyrArgGlyLysGluTyrGlyInThrMetThrValSerAsnProGlnGlyCysArgLeu 156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 731 GTGGGCTCCGAAGTGGGA-----GACAGGACGCTGCTGGATGGCCACTCACACTG 781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 157 PheTyrGlyAspLeuGlyProMetProAspGlnGluLeuPheGlyProVal**Leu 176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 782 CCAGAC-----CTGGCAGTCC---CTGACAGACAGGGGAGTGAAGTACAGT 829
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 830 AGGCATGCTGAGCTGGTGGGAGCTGCTCTGCGGGCGGGCGGCGAGTGGCTC 889
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 SerLysLeuLeuAspValMetAspArgGlyLeuLeuLeuValSerGlyHisAlaIle 216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 890 TGGGCGCCAGCGCTGGGCGCTGCCACACATATCTGGCA-----GTGAGCGAGG 940
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 TyrAlaIleArgLeuCysGlnCysLysValTyrTrpSerGlyProCysAlaProSerLeu 236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 941 CTGCTCCCAACAGCGGGGATGGCTGATGCGAGTCCCAAGGACAAGGAAGGAGGC 1000
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 ValAlaProAsnLeu-----IleGluArgGlnLysLysValLys 249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1001 GTTTTGACCTGGGCGCTTCTATTGATCTGATCTTACCGAAGGAGC---CGA 1057
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 LeuPheCysLeuGluThrPheLeuSerAspLeuIleAlaHisGlnLysGlyGlnIleGlu 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1058 CGCTCACACGCTATGCTCTGCTGTGTGGGGAGTCAATGCGCCCGAGGAGCGCG 1117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 LysGlnProProPheGluIleTyrLeuCysPheGlyGluGluTrpProAspGlyLysPro 289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1118 TGGACCAAGAGCTCGTGAAGTCAAGTGTGCGCCAGTGCCTCAGGCGCTGTGAGAA 1177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 LeuGluArgLysLeuIleLeuValGlnValIleProValValAlaArgMetIleTyrGlu 309
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1178 ATGGCCCGGTAGGGGTGCTCTCCCTCGTGGAGAT---ACTGTGAGCTGCACATTTC 1234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 MetPheSerGlyAspPheThrArgSerPheAspSerGlySerValArgLeuGlnIleSer 329
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1235 AACAGCCACCCACTCTCCCTCACCTCCGACAGTACAAGGCTACTCGAGGACTTGTG 1294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 330 ThrProAsp-----IleLysAspAsnIleValAlaGlnLeuLysGlnLeuTyr 345
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1295 GAGGCATGGATTCAGGCGCTCGGAGAGCTGAGCCCTGCTCTCATGTTGTGCT 1354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 ArgIleLeuGlnThrGlnGluSerTrpGln----- 355
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1355 CCAACCCCTGTTCCCGACACCTCAACCAATAAAGTGTCTCTGCTATG 1405
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 356 -----ProMetGlnProThrProSerMet---GlnLeuProProAlaLeu 369
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-882-171-374
; Sequence 374, Application US/09882171
; Publication No. US20030175858A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
```

[illegible]

```

QY 674 TACCGGGCCGCAAGTCTTCCAG---CAGACCATCTCTGCGCCGAGGCGCTGGCGTG 730
Db 137 TyrArgGlyLysGluThrValSerAsnProGlnGlyCysArgLeu 156
QY 731 GTGGGTCGCAAGTGGGA-----GACAGGACGCTGCTGGATGGCCAGTCACACTG 781
Db 157 PheTyrGlyAspLeuGlyProMetProAspGlnGluLeuPheGlyProVal***Leu 176
QY 782 CCAGAC-----CCTGGCATGTC---CTGACAGACAGGGAGTGATGAGTACAGTG 829
Db 177 GluGlnValLysPheProGlyProGluHisIleThrAsnGluLysGlnLysLeuPheThr 196
QY 830 AGGCATGTCTGAGTCTGCTGGTGGGAGCTCTCTCTGCGGGCCGCGGAGTGGCTC 889
Db 197 SerLysLeuLeuAspValMetAspArgGlyLeuLeuLeuValSerGlyHisAlaIle 216
QY 890 TGGGCCACGCGCTGGGCACTGCCACATACTGGCA-----GTGAGCGAGGAG 940
Db 217 TyrAlaIleArgLeuCysGlnCysLysValTyrTrpSerGlyProCysAlaProSerLeu 236
QY 941 CTGCTCCCAACAGCGGCGATGGCGCTGATGGAGGTCCCAAGACAAAGGAAGAGGC 1000
Db 237 ValAlaProAsnLeu-----IleGluArgGlnLysLysValLys 249
QY 1001 GTGTTTGACCTGGGCGCTTCACTGTAGATCTGATTACCTTCCAGGAAGAGC---GGA 1057
Db 250 LeuPheCysLeuGluThrPheLeuSerAspLeuIleAlaHisGlnLysGlyGlnLeu 269
QY 1058 CGCTCACCAGCGTATGCCCTCTGTTGTGGGGAGTCTATGGCCCCCAGGACGAGCG 1117
Db 270 LysGlnProProGluIleTyrLeuCysPheGlyGluGluTyrProAspGlyLysPro 289
QY 1118 TGGACCAAGAGCTGCTGATGTGAAGTTGTGGCCACGTGCTCAGGCGCTGTGTAGAA 1177
Db 290 LeuGluArgLysLeuIleLeuValGlnValIleProValValAlaArgMetIleTyrGlu 309
QY 1178 ATGGCCCGGGTAGGGGGTCCCTCTCCCTCGGAGAT---ACTGTGAGCTGACATTTCC 1234
Db 310 MetPheSerGlyAspPheThrArgSerPheSerGlySerValArgLeuGlnIleSer 329
QY 1235 AACAGCCACCCACTCTCCCTCACCTCCGACCAAGTACAAAGGCGCTACTCGAGACTGTGTG 1294
Db 330 ThrProAsp-----IleLysAspAsnIleValAlaGlnLeuLysGlnLeuTyr 345
QY 1295 GAGGCGATGGATTCCAGGCGCTGGGAGAGCTGAGCGCTGCTCTCATGTGTGCT 1354
Db 346 ArgIleLeuGlnThrGlnGluSerTrpGln----- 355
QY 1355 CCAACCCCGCTGTTCCCAACCACTCAACCAATAAAGTGTCTCTGCTATG 1405
Db 356 -----ProMetGlnProThrProSerMet---GlnLeuProProAlaLeu 369

RESULT 9
US-10-164-861-374
; Sequence 374, Application US/10164861
; Publication No. US20030225248A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT FILING DATE: 2002-06-10
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
US-10-164-861-374

```

```

Alignment Scores:
Pred. No.: 6,93e-08 Length: 373
Score: 246.00 Matches: 104
Percent Similarity: 42.32% Conservative: 64
Best Local Similarity: 26.20% Mismatches: 161
Query Match: 9.11% Indels: 68
DB: 14 Gaps: 20

US-09-975-253-1 (1-1426) x US-10-164-861-374 (1-373)
QY 329 GACCGAGCAAGGAC---CCTCAGCAGCCACATAAAATCTACGAGTTTGTGAATCAGGA 385
Db 3 AspGlyThrLysGluValProMetAsnProValLysIleTyrGln----- 17
QY 386 GTTGGGGACTTTTCCAGCCAGACACCTCT-----CCGACACCAATCGTGGAGGC 436
Db 18 ValCysAspIleProGlnProGlnGlySerIleIleAsnProGlySerThrGlySerAla 37
QY 437 AGTACTTCTGAT-----ACCCAGGAGAGACATCTCTGATGAGTACTTGGGT 481
Db 38 ProTrpAspGluLysAspAsnAspValAspGluGluAspGluLysLeuAspGln 57
QY 482 AACATGTTGTGGCCCCACTCCAGATCCGGGACCCCAAGCCTGGCTGATGCCCTCGAG 541
Db 58 SerGlnHisHisValProIleGlnAspThrPhe---PropheLeuAsnIleAsnGlySer 76
QY 542 CCC---TGCCCTCAGCCCTCGGAGCCCGAGTGGACAATCCCACTCCC----- 589
Db 77 ProMetAlaProAlaSerValGlyAsnCysSerValGlyAsnCysSerProGluAlaVal 96
QY 590 TTCCCAACACCTGGGCGCTCTGAG-----AACCCACTAAGCGGCTGTG 634
Db 97 TrpProLysThrGluProLeuGluMetGluValProGlnAlaProIleGlnProPheTyr 116
QY 635 GTGCGCGGGGAAGAGTGG---GAGTTCCGAGGTGACAGCCCTC 673
Db 117 SerSerProGluLeuTyrIleSerSerLeuProMetThrAspLeuAspIleLysPheGln 136
QY 674 TACCGGGCCGCAAGTCTTCCAG---CAGACCATCTCTGCGCCGAGGCGCTGGCGTG 730
Db 137 TyrArgGlyLysGluTyrGlyGlnThrMetThrValSerAsnProGlnGlyCysArgLeu 156
QY 731 GTGGGTCGCAAGTGGGA-----GACAGGACGCTGCTGGATGGCCAGTCACACTG 781
Db 157 PheTyrGlyAspLeuGlyProMetProAspGlnGluLeuPheGlyProVal***Leu 176
QY 782 CCAGAC-----CCTGGCATGTC---CTGACAGACAGGGAGTGATGAGTACAGTG 829
Db 177 GluGlnValLysPheProGlyProGluHisIleThrAsnGluLysGlnLysLeuPheThr 196
QY 830 AGGCATGTCTGAGTCTGCTGGTGGGAGCTCTCTCTGCGGGCCGCGGAGTGGCTC 889
Db 197 SerLysLeuLeuAspValMetAspArgGlyLeuLeuLeuValSerGlyHisAlaIle 216
QY 890 TGGGCCACGCGCTGGGCGACTGCCACATACTGGCA-----GTGAGCGAGGAG 940
Db 217 TyrAlaIleArgLeuCysGlnCysLysValTyrTrpSerGlyProCysAlaProSerLeu 236
QY 941 CTGCTCCCAACAGCGGCGATGGCGCTGATGGAGGTCCCAAGACAAAGGAAGAGGC 1000
Db 237 ValAlaProAsnLeu-----IleGluArgGlnLysLysValLys 249
QY 1001 GTGTTTGACCTGGGCGCTTCACTGTAGATCTGATTACCTTCCAGGAAGAGC---GGA 1057

```

Db	250	LeuPheCysLeuGluThrPheLeuSerAspLeuIleAlaHisGlnLysGlyGlnIleGlu	269
Qy	1058	CGCTCACACGCTATGCCCTCTGCTGTGTGGGGAGTCAATGCCCGCCAGGACGAGCGG	1117
Db	270	LysGlnProPheGluIleLeuValGlnValIleProValAlaAlaArgMetIleIleGlu	309
Qy	1118	TGGACCAAGAGCGCTGATGCTCAAGTGTGTGGCCCGCTCCCTCAGGCGCTTGTAGAA	1177
Db	290	LeuGluArgLysLeuIleValGlnValIleProValAlaAlaArgMetIleIleGlu	309
Qy	1178	ATGGCCCGGTAGGGGTGCTCTCCCTGAGAAAT---ACTGTGACCTGCACATTTCC	1234
Db	310	MetPheSerGlyAspPheThrArgSerPheAspSerGlySerValArgLeuGlnIleSer	329
Qy	1235	AACGCCACCCACTCTCCCTCACTCCGACAGTACAAGGCTCACTGCAGGACTTGTGTG	1294
Db	330	ThrProAsp-----IleLysAspAsnIleValAlaGlnLeuLysGlnLeuTyr	345
Qy	1295	GAGGCGATGGAFTTCCAGGGCCCTGGGAGAGCTGAGCCCTCGCTCTCATGTGTGCT	1354
Db	346	ArgIleLeuGlnThrGlnLysSerTrpGln-----	355
Qy	1355	CCAAACCCCTCTTCCCAACCACTCAACCAATAAAGTGTCTCTGCTATG	1405
Db	356	-----ProMetGlnProThrProSerMet-----GlnLeuProProAlaLeu	369

RESULT 10			
US-10-084-846A-7			
; Sequence 7, Application US/10084846A			
; Publication NO. US2004006026A1			
; GENERAL INFORMATION:			
; APPLICANT: WEITNAUER, GABRIELE			
; APPLICANT: MUHLNBERG, AGNES			
; APPLICANT: TREPZER, AXEL			
; APPLICANT: BECHTHOLD, ANDREAS			
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES			
; FILE REFERENCE: 1974-005			
; CURRENT APPLICATION NUMBER: US/10/084,846A			
; CURRENT FILING DATE: 2003-02-25			
; PRIOR APPLICATION NUMBER: PCT/EP01/09815			
; PRIOR FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: DE 101 09 166.4			
; PRIOR FILING DATE: 2001-02-25			
; NUMBER OF SEQ ID NOS: 120			
; SOFTWARE: PatentIn Ver. 3.2			
; SEQ ID NO 7			
; LENGTH: 19652			
; TYPE: PRT			
; ORGANISM: Streptomyces viridochromogenes			
; FEATURE:			
; OTHER INFORMATION: Protein 2: amino acid sequence encoded by coding strand 2.			
; OTHER INFORMATION: Start codon: gat, Start position: nucleotide 2.			
US-10-084-846A-7			

Alignment Scores:			
Pred. No.:	1.74e-07	Length:	19652
Score:	243.50	Matches:	138
Percent Similarity:	31.80%	Conservative:	28
Best Local Similarity:	26.44%	Mismatches:	183
Query Match:	9.10%	Indels:	173
DB:	15	Gaps:	27

US-09-975-253-1 (1-1426) x US-10-084-846A-7 (1-19652)			
Qy	1382	TTGAGGTGGGGGACAGGGGGGTGGAGGCACACCATGAGGAGCGAGGCTCAGCTCT	1323
Db	2202	ValArgArgTrpProGlySerSerCysArgAlaProSerArgArgGlySerSerGly	2221
Qy	1322	-----CCCCAGGGC	1314
Db	2222	AlaGlyArgGlyArgSerGlyCysArgArgProGlySerArgAlaValArgProGlyGly	2241

Qy	1313	-----CCTGGAATTCATGCCCTCCACCAAGTCTCTGAGGT-----AGG	1275
Db	2242	AlaCysArgArgAlaProGlySerProAlaArgProArgProAlaGlyThrProArg	2261
Qy	1274	CCTTGTACTGTCGAGGTGAGGG-----AGAGTGGGTGGCTGTGGAAATGT	1227
Db	2262	ProProArgGlyArgGlyArgGlyProArgAlaArgAlaGlyArgSerThrArgArg	2281
Qy	1226	GCAGTCCACAGTATTCTCCAGGAGGAGGACCCCTTACCCTGAGGAGGCTTCTACCAAGG	1167
Db	2282	SerAlaAlaArgSerArgProGlyArgProArgSerProGly-----	2296
Qy	1166	CCCTGAGGACGTCGGGCACAACTTGACATCAACAGAGCTCTTGTGTCACGGCTGCTCT	1107
Db	2297	-----GlyAlaAlaGlyArgProArgProGlyArgAlaAlaValAla-----	2310
Qy	1106	GGGGCCATGACTCCCCACACAGAACAGAGGCATAGCGTGGTGGAGCTCGCTCTCTT	1047
Db	2311	-----HisArg-----ArgGlyArgProAlaArgSerValArgGlySer	2323
Qy	1046	CGTGAAGGTAAATCAGATCTACAAATGAAGGCGCCAGGTCAAAACACGCTCTCTCT	990
Db	2324	ProProArgProPro-----AlaProValProValArgProAlaProArg	2338
Qy	989	-----TGTCCTTGGGGACCTCCGCATGTGTGGCAGTGGCCCGCCAGCGCTG	888
Db	2339	ThrAlaCysAlaAlaGlyLeuProAlaProProArgProAlaAlaGlyArgAlaSer	2358
Qy	947	GGAGCAGCTCTGCTCACTGCCCGAGTGTGTGGCAGTGGCCCGCCAGCGCTG	888
Db	2359	SerSerAlaProArg-----ProProArgArgProGlyValPro-----ArgGlyProPro	2375
Qy	887	GCCACTGCCCGCGCCGACAGAGCCAGTCCCGCCAGGAGGAGTCAACATGCTCTCA	828
Db	2376	ProProAlaArgAlaAlaArgProValArgTrpSerGlyAlaAlaGlyCysProArg	2395
Qy	827	CGTAGCTCATCACTCCCTCTCTCAGGACATGCGAGGCTGTCAGTGTGCTGCTG	768
Db	2396	Arg-----CysVal-----	2398
Qy	767	ATCCAGGAGCGTCTCTGCTCTCCACTTCGACCCACACAGCGCAGGC-----CCTCG	714
Db	2399	-----ArgArgCysCysProValGlnProArgArgSerAlaGlyArgProPro	2415
Qy	713	GGCAGGAGATGTCTCTCGAAGACTTGGCGGCGCCCGGT-----AGAAGG	669
Db	2416	SerGlyArgArgSerAlaAlaArgProArgGlyAlaAlaGlyThrSerArgArg	2435
Qy	668	CTGTCACTCGAACTCCACT-----CTTCCCGCCGACCAACAGCGCTTCA	621
Db	2436	ArgAlaProGlyArgProSerGlyThrArgProSerProProGlyAlaAla-----	2454
Qy	620	GTGGGTCTCAGAGGCGCCAGGTGGGAGGAGTGGGATGTGCAAGCTGGGCTCC	561
Db	2455	-----CysProArgGlyProThr	2460
Qy	560	GCAGGG-----GCTGAGGCGAGGCTCAGGGG	534
Db	2461	AlaGlyProProGlyProProAlaArgAlaGlyProValAlaHisGlySerValProGly	2480
Qy	533	-----CTACAGCAGGCTTGGGGTCCCGATCTGGAGTGGGGCAACACCATGTTAC	480
Db	2481	ProProValArgProArgArgValArgAspArgProProAlaProProArg	2500
Qy	479	CCAGTAATCATCA-----	465
Db	2501	GluAlaAlaArgProGlyProProProProAlaHisSerAlaProSerArgProArgPro	2520
Qy	464	-----GAATGTCTTCTGGGTATCAGAAGTACTGCTCCACCATGG	423
Db	2521	SerHisArgValProAlaSerCysHisProGlyGlyArg-----IleAspGlyTrp	2538
Qy	422	-----TGTCCGGAGAGGTGCTGGCTGGGAAAGTCCCAACTCTCTGAGTTCACAA	372


```

Db      2539 CysAlaGlyCysProAlaThrLeuThrIleGlyArg-:::
QY      371 ACTCGTAGATTTATGTGGTCTGTAGGTCCTTGTCTCGGTCCTCTGCTAAACGCAACC 312
Db      2554 ThrCysProLeuThrValGlySerArgGlyAlaLeu-:::
QY      311 CTTCTTTGGGTTGAGGCGAGCGGAAATCTCTCTCCAGGTGGCAGGTCTGGCTTAT 252
Db      2566 -----GlyGlnAlaAspArgLeuLysAlaGluIleGlyArgLeuAengly 2580
QY      251 CCTCCCGGAAACATATGCACCATGTCCTCGCCAGCGCTGGAAGATTCGGAATCCT 192
Db      2581 ProValArg-----SerTrpProAlaProAlaValProAspThrArgArgSer 2596
QY      191 CTGCTGTGCATCTCCGCTA-----GGCGGTGCTTCCAGGATGCGGAAGCGGTGC 138
Db      2597 ArgAlaAlaSerProThrThrGlyProGlyArgAlaGlyArgAspArgGlyProGlyArg 2616
QY      137 GGCTCTGTTCACCCAGGCGACGCTCCAGTTCGCTCCAGGTCGACCTCGGACACGAGCC 78
Db      2617 GlyAlaArgSerValProSerArgProProVal-----ProAlaTrpArgArgCys 2633
QY      77 AGGGCA 72
Db      2634 ArgAla 2635

```

```

RESULT 11
US-09-805-020-46
; Sequence 46, Application US/09805020
; Publication No. US20020086384A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-46

```

```

Alignment Scores:
Pred. No.: 1.67e-07 Length: 292
Score: 240.00 Matches: 65
Percent Similarity: 43.66% Conservativity: 28
Best Local Similarity: 30.52% Mismatches: 85
Query Match: 8.89% Indels: 35
DB: 9 Gaps: 6

US-09-975-253-1 (1-1426) x US-09-805-020-46 (1-292)

```

```

QY      47 ATGGGAACCCCAAGCCACGNTCTGCTCGGTGTGTGTCGACGTGGAGCTGGGGCAA 106
Db      1 MetProIleThrArgMetArgMetArgProIlePueGluMetGlnIleAsnSerAsnGln 20
QY      107 CTGGAGGCGTGGCTGGTGAACAAGACGCGCTTCGCGATCCCTTGAAGCAC 166
Db      21 IleProGlyLeuIleTrpIleAsnLysGluGluMetIlePheGlnIleProTrpLysHis 40
QY      167 GGCCTACGCGAGGTGCACAG---CAGGAGATTTCCGAATCTTCAGGCTGGGCCGAG 223
Db      41 AlaAlaLysHisGlyTrpAspIleAsnLysAspAlaCysLeuPheArgSerTrpAlaIle 60
QY      224 GCCACTGTGTGATATGTTCCCGGAGGAGGATAGCCAGACCTGCCAAGCTGGAAGGAT 283
Db      61 HisThrGlyArgTyrLysAlaGlyGluLysGluProAspProLysThrTrpLysAlaAsn 80
QY      284 TTCGCTCTGCCTCAACCCGCAAGAGGGTGTGGTTTAGCAGAGGACCGGACGACAGGAC 343

```

```

Db      81 PheArgCysAlaMetAsnSerLeuProAspIleGluGluValLysAspGlnSerArgAsn 100
QY      344 CCT---CAGCACCACATAAATCTACGAGTTGTG-----:::
Db      101 LysGlySerSerAlaValArgValTyrArgMetLeuProLeuThrLysAsnGlnArg 120
QY      377 -----AACTCAGGAGTTGGGACTTTTCCAG 403
Db      121 LysGluArgLysSerLysSerArgAspAlaLysSerLysAlaLysArgLysSerCys 140
QY      404 CCAGACACCTCTCCGAGACCAATGTGTGAGGACGACTCTCTCATACCCAGGAGACATT 463
Db      141 GlyAspSerSerProAspThrPheSerAspGlyLeuSerSerThrLeuProAsp-As 160
QY      464 CTGGATGATTTACTGTGTAACATGTGTGTGGCCCACTCCAGATCCGGGACCCCAAGC 523
Db      160 phisSerSerTyrThrValProGlyTyrMetGlnAspLeuGluValGln----- 177
QY      524 CTGGCTGTAGCCCTGAGCCCTGCTCAGCCCTCAGCCCTCGGAGCC-----CCAGC 571
Db      178 -----AlaLeuThrProAlaLeuSerProCysAlaValSerSerThrLeuProAs 194
QY      572 TTGGACAATCCCA-----CTCCTTCCCAACC 599
Db      194 ptrHisIleProValGluValProAspSerThr 206

```

```

RESULT 12
US-10-084-846A-6
; Sequence 6, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 19662
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-6

```

```

Alignment Scores:
Pred. No.: 3.2e-07 Length: 19662
Score: 239.50 Matches: 157
Percent Similarity: 32.45% Conservativity: 27
Best Local Similarity: 27.69% Mismatches: 209
Query Match: 8.95% Indels: 174
DB: 15 Gaps: 33

US-09-975-253-1 (1-1426) x US-10-084-846A-6 (1-19662)

```

```

QY      1379 AGTGGTGG-----GGAACAGGGGGTTGAG 1353
Db      12610 ArgTrpTrpAlaAlaSerArgCysTrpProAlaArgProGlyArgGly----- 12626
QY      1352 GCACCATGAGGAGCGAGGCTCAGCTCTCCCGAGGCCCTGGAATCCATGCCCTCCA 1293
Db      12627 -----AlaSerProAlaGlyGlyProProProSerSer 12638
QY      1292 CCAAGTCTCGAGTAGGCTTGTACTGTGTCGGAGGTGAGGGAGTGGGTGTGTGG 1233

```

12961 -----GlySerArgArgArg-----GlyArgAlaArgProLeu 12971
 335 TCCGGTCTCTGCTAAACGCAACCTTCTTTGCGGTTGAGGCGAGAGCGGAAATTCCTCT 276
 12972 AlaProArgArgTrpAsnAlaThr-----AlaGlyGlyAlaAlaArg 12985
 275 TCCAGGTTGGCAGGCTCG-----GCTTATCCCTCC 246
 12986 GlyArgAlaAlaGlyLeuArgGlySerArgSerAlaGlyAlaProThrSerProArgSer 13005
 245 CGGAAACATATGCACCATGCTCCG-----CCAGGCTCTGGA--- 207
 13006 ArgArgGlySerGlyArgArgProArgArgArgGlyArgSerGlyProArgAlaGlyCys 13025
 206 -----AGATTCCGAAATCTCTCTGCTGTCATCTCTGCGTAGGCGCT 165
 13026 AlaValAlaSerValArgArgThrArgProArgArgValArgProAlaSerArgGly 13045
 164 GCTTCC-----AAGGATGCGGAAGCGGCTCGGCTCTTGTTCACCCAGGCGCAGC 114
 13046 SerSerAlaArgProArgGlyAlaGlyThrAlaArgArgThrProGlyGlyArgProArg 13065
 113 CCTCCA-----GTTCCCGCAGGTCCAGCTGCGACACACCCAGGCGG--- 74
 13066 ProProArgSerArgArgArgAlaAlaProGlyCysAlaGlySerPro-ProSerThrAr 13085
 73 -----CAGGANCCGCTGCTTTGGGGTTCCCATGCTCGGCTTACGATGGAAGTCCGG 21
 13085 gArgArgArgSerProTrpProAlaGlyProAlaGlyProArgAlaProSerCys-----ArgAlaAl 13103
 20 GCGTCCGCGGACGTGGAC 2
 13103 aCysAlaValArgTrpAsp 13109

RESULT 13

US-10-084-846A-8
 ; Sequence 8, Application US/10084846A
 ; Publication No. US20040006026A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEITNAUER, GABRIELE
 ; APPLICANT: MUEHLENWEG, AGNES
 ; APPLICANT: TREFFER, AXEL
 ; APPLICANT: BECHTHOLD, ANDREAS
 ; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
 ; FILE REFERENCE: 1974-005
 ; CURRENT APPLICATION NUMBER: US/10/084,846A
 ; CURRENT FILING DATE: 2003-02-25
 ; PRIOR APPLICATION NUMBER: PCT/EP01/09815
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: DE 101 09 166.4
 ; PRIOR FILING DATE: 2001-02-25
 ; NUMBER OF SEQ ID NOS: 120
 ; SOFTWARE: Patent in Ver. 3.2
 ; SEQ ID NO 8
 ; LENGTH: 19608
 ; TYPE: PRT
 ; ORGANISM: Streptomyces viridochromogenes
 ; FEATURE:
 ; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
 ; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
 US-10-084-846A-8

Alignment Scores:
 Pred. No.: 6,83e-07 Length: 19608
 Score: 234.50 Matches: 153
 Percent Similarity: 33.16% Conservative: 39
 Best Local Similarity: 26.42% Mismatches: 208
 Query Match: 8.78% Indels: 179
 DB: 15 Gaps: 32

US-09-975-253-1 (1-1426) x US-10-084-846A-8 (1-19608)

QY 1379 AGTGGTGGGGAACAGGGGGTTGGAGGCACACCATGAGGAGCGGCTCAGCTCTCC 1320

12639 ProSerArgAlaProAlaProAlaGly-----GlyGluArgAlaGlyProTrp 12655
 1232 -----AAATGTGCAGGTCCACAGTATTCTCCAGGAGGAGG 1197
 12656 GlyArgValAlaSerProAlaArgSerAlaArgAlaProArgArgAlaProGly----- 12673
 1196 CACCCCTTACCGCGGCCATTCTACCAAGGCCCTGAGGACGCTGGCGCACACCTTGACCA 1137
 12674 AsnAlaAlaProAlaThrAspProArg-----ThrThrProArgProArgPro 12690
 1136 TCACGA-----GCTCTGTGTCCACGCTGGTCTGGGGCCATGACTCCCGCCACACAGA 1083
 12691 ArgArgArgProAlaArgCysProAlaArgProAlaAlaArgProProGlySerAla 12710
 1082 ACCGAGGCGATAGCGTGTGAGCTCGCTTCCTTCGCTGAAGTAATCAGATCTACAA 1023
 12711 GlyArgAlaGlySerProProGlyVal-----Pro 12720
 1022 TGAAGGCCCCAGGTCAACA-----CGCTCTCTCTCTGTCT 984
 12721 ProArgProProGlyArgAlaArgCysArgArgArgGlyAspArgAlaAlaArgProArgPro 12740
 983 TGGGACCT-----CGCCATCAGGCCCATGCGCGCTGTTGGGAGCA 942
 12741 SerProProProValAlaCysArgArgArgSerArgArgAlaAlaArgProCysGlyAla 12760
 941 GCT-----CCTCGCTCACTGCCAGTATGTGGCAGT-----GCCCGACCGCTGGCGCC 891
 12761 AlaGlyGlyProArgArgProAlaArgSerThrSerArgArgProAlaAlaGlyPro 12780
 890 AGAGCCACTCCCGCGCCCGCCAGAGAGCA-----GTCCCCCAGCCAGCGCTCAGCA 837
 12781 Arg---ThrArgArgAlaAlaAlaProGluArgAlaValHisArgAspArgArgThr 12799
 836 CATGCTCAGTACTCA-----819
 12800 ValAlaArgSerAsnArgAsnArgValArgProProArgProAlaAlaGly 12819
 818 -----TCACTCCCTCTCTGTCAGGACATGCCAGGTCTGGCAGTGTGACTG 771
 12820 ArgProGlyArgGlyArgProValArgAlaGlyAlaArgProGly-----12834
 770 GCCATCAGCAGCTCTCTCCACTTCGACCCCGCCAGCCAGCGGCTCCCGGC 711
 12835 -----ProAlaAlaThrAlaProProAspArgArgGlyProAlaArg---ProSerGly 12851
 710 AGGAGTGGTCTGCTGGAAGACTTGGCGGCCCGGTAGAGGCTGTCCACCTCGA---ACT 654
 12852 ProArgHisGlyArgAlaAlaArgGlySerGlyArgArgArgProArgProSerGlyArg 12871
 653 CCCACTCTTCCCGCGCACCAACAGCCGCTTCACTGGGTCTCAGAGGGCCCGAGGTTG 594
 12872 ProAlaAlaProArgProArgProArgGlyArgProGlySerAlaAlaAspProProSerAla 12891
 593 -----GGAAGGAGTGGGATTTCCAGCTGGGCTCCCGAGGCT-----552
 12892 ProGlyGlySerSerCysPro-----GlyGlyArgGlyAlaLeuArgProArgArg 12909
 551 -----GAGGGCAGGCTCAGGGGCTACAGCCA-----GGCTTGGG 516
 12910 ArgProProGlyThrProProAlaArgArgProGlyArgSerProArgArgGlyArgGly 12929
 515 GTCCCGGATCTGGAGTGGGCGCCACACCATGTTACCGAGTAATCTATCCAGATGCTT 456
 12930 CysProGlyThrGlyGlyCysArgArgProAlaCysCysAlaThrAlaAlaGlyArgCys 12949
 455 CTGGGTATCAGAGTACTGCTCCACCATTTGGTCCGAGAGAGTGTCTGCTGGGAAA 396
 12950 ProArgArgArgArgAla-----CysProAlaAlaCys-----12960
 395 AGTCCCCCACTCTCAGTTCAAACTCGTAGATTTTAATGGTGTGGTGGTCTTTC 336

Db 11090 ArgArgTrrAlaArgProGly---ArgProArgThrArgArgAlaAlaProValPro 11108
 QY 1319 CAGGGCCCTGGAATCCATCCCTCCACCAAGTCTCGAGGTAGCCCTTGTACTGTGTCGG 1260
 Db 11109 ArgGlyAlaAlaGlyProValProGlyProSerProAlaSerProSerCysAlaArgGly 11128
 QY 1259 AGGTGAGGAGCA---GTGGTGGCTTGGAAATGTCA 1224
 Db 11129 Arg---GlyArgAlaHisArgProArgGlySerSerSerArgAla 11147
 QY 1223 GTTCCACAGTATTCT---CCAGGGAGG 1200
 Db 11148 GlyProArgArgSerGlyAlaValArgArgProProAspGlyArgSerArgProGlyArg 11167
 QY 1199 AGCACCCCTACCCGGCCA---TTTCTACCAAGCCCTGAGCAGCGTGGGCACACACT 1143
 Db 11168 Glu---ProGlyProArgSerArgProValGlyThrAlaAlaArgSerPro 11184
 QY 1142 TGACCATCAGAGCCTCT---TGGTCCACGGCTGCTCTGGGCCATGACTCCCCACAC 1086
 Db 11185 ArgProGlyProValSerProTrpSerTrp---MetArgProProPhe 11199
 QY 1085 AGAACACAGAGGATAGCGTGGTGGAGCTCCGCTTCTCCGTGAAGTAATCAGATCTA 1026
 Db 11200 AlaserSerGlyGln---ArgHisGluArgGlyThrProAlaLeuArg---GlnAla 11217
 QY 1025 CAAATGAAGGCCCCAGAGTCAACACACCGCTCTTCTTGTCTTGGGACCTCCGCATCAG 966
 Db 11217 laThrMetSerGlnThrSerAlaValAlaArgSerCysSerThrSerValLysP 11237
 QY 965 GCCATGCGCGTGTGGGAGCAGCTCTCGCTCACTGCCCGCATGTG--- 916
 Db 11237 roAlaCysSerArgAlaAlaArgSerAspGlyArgSerGlnArgThrSerThrLeuA 11257
 QY 915 ---TGACGTGCCCGCCAGCCGCTGG---GCCAGAGCCACT 882
 Db 11257 rgProArgAlaTrpGlySerGlyTrpSerSerArgThrGlyTrpAlaAlaTyArgArg- 11276
 QY 881 GCCCGCCCGCAGAGCAGCAGTCCCGCCAGCAGCAGCTCAGCATCCCTCAGCTAGC 822
 Db 11277 GlyArgSerPheArgProSerCysProSerSerGlyTrpArgGlyArgGlySerArg 11296
 QY 821 TCATCACTCCCTGTCTGTCAGGACATCCAGGCTCTGGCAGTGTGCTGCCATCCAG 762
 Db 11297 ArgSer---ArgCysArgA-gAlaAlaProArgAlaAlaAlaArg 11310
 QY 761 GCAGCGTCTGTCTC---CCACTCGGACCCCA 729
 Db 11311 GlyArgGlyAlaLeuArgAlaTrpArgArgSerGlyArgAlaArgArgThrProSer 11330
 QY 728 GCC---GCAGGC---CCTCCGGCAGGAGATGG 702
 Db 11331 ThrProArgArgAlaGlyValProArgTrpArgAlaProProArgThrArgArg 11350
 QY 701 TCTGTGGAAGACTTGGCGGCC---GGTAGAGGCTGTCACTCGA 657
 Db 11351 GlyAlaGlySerLeuArgGlyProProGlyArgGlySerArgArgAlaValProThr 11370
 QY 656 ACTCCACTCTTCCCGCCGACCAACAGCCGCTTTCAGTGGGTCTCAGAGGCGCCCGAGT 597
 Db 11371 SerProGlyCysProArgProPro---AlaArgAlaPro--- 11383
 QY 596 TTGGAGGGAGTGGATTGTCAGAGTGGGCTCCGAGGGGTGAGGCGAGGCTCAG 537
 Db 11384 ---GlyArgAlaTrp---Pro-ProGlyThrProArgArgThrArgArgProAlaPr 11400
 QY 536 GGGCTACAGCAGGCTTGGGGTCCCGATCTGG---CAGTGG--- 497
 Db 11400 oGlyAlaAlaGlySerAlaGlySerArgAlaTrpArgArgTrpArgThrGlyProArgC 11420
 QY 496 ---GGCCAACACATGTT---ACCCA 477

Db 11420 sThrGlnSerArgSerProThrAlaArgValHisValProAspAlaGluSerGlnAspAr 11440
 QY 476 GTAACCTCATCAGAAATGCTTCTCTGGTATCAGAGTACTGCCTCCACCATTTGTGTCCG 417
 Db 11440 sSerLeuLeuProTyrlzHisCysGlnValArgPheGlnAlaArgArgLeuGlyLeu-G 11460
 QY 416 GAGAGGTGTCTGGCTGGGAAAGTCC---CCAACTCCTCAGT 378
 Db 11460 ly---lleThrGlyArgArgSerAlaGlyArgProHisThrAlaProThrPro--- 11477
 QY 377 TCACAAACTCGTAGATTTTATGTGGTCTGTGAGGGTCTTGTCTCCCGTCTCTCTCTAAAC 318
 Db 11478 ---GlyArgAlaLeuArgArgSerArgSerA 11487
 QY 317 GCAACCTTCTTTCGGTTCAGGCGAGAGCGGAAATCTCTTCCAGGTGGCAGGTCTG 258
 Db 11487 rgArgSerAlaValArgGlyArgArg-ArgProThrArgArgSerGlyArg 11506
 QY 257 GCTTAT---CCTCCCGGGAACATATGCAC---CAGTGG 225
 Db 11507 AlaArgProAlaGlyCysProProSerArgArgArgProArgGlyProProArgTrp 11526
 QY 224 ---CCTCGCCCGCTGGAAGATTC--- 201
 Db 11527 ArgCysArgArgProSerProGlyArgAlaGlyAlaArgTrpGlyGlyTrpGly 11546
 QY 200 CGAAATCTCTCTGTGTCATCTCGTAGCCGCTTCCAGGGATCCGGAAGCGG 141
 Db 11547 HisArgProProHisAlaArgProGlyProAlaProAlaArgCysTrpCysGlyThrSer 11566
 QY 140 TCGGCTCTTGTTCACCCAGGCGCCCTCCAGTTGCCCGAGT--- 96
 Db 11567 ArgArgArgProGlyProArgProSerProSerArgGlyArgAlaArgArgProArgCys 11586
 QY 95 CCAGTTCGACACACAGC---AGGGCA 72
 Db 11587 ProArgSerThrProAlaProAspArgArgArgArgProArgSerAlaHisArgAla 11606
 QY 71 GGANCCGTGCTTGGGGTTCCTCCATGGTCCGCTTCCAGTACGATGGAAG 27
 Db 11607 Gly---SerGlyArgGlyTrpArg 11613

RESULT 14

US-10-755-889-356
 ; Sequence 356, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
 ; FILE REFERENCE: D0284 NP
 ; CURRENT APPLICATION NUMBER: US/10/755,889
 ; CURRENT FILING DATE: 2004-01-13
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068
 ; PRIOR FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757
 ; PRIOR FILING DATE: 2003-05-12
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 356
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-755-889-356

Alignment Scores:

Pred. No.: 5,59e-07 Length: 164
 Score: 231.50 Matches: 47
 Percent Similarity: 55.75% Conservative: 16
 Best Local Similarity: 41.59% Mismatches: 39
 Query Match: 8.58% Indels: 11
 DB: 16 Gaps: 3

1322	QY	CCCAGGCGCCCTGGAAATCCA-----	1302
369	Db	ProHisGlyProGlyGlyProGlnProLeuAlaSerLeuPheTrpMetPhe	388
1301	QY	TGCCCTCCACCAAGTCTCGAGTAGG-----CCTGTACTGGT	1263
389	Db	CysGlnProProProHisProGlnPheLeuTrpHisArgProLeuValThrGly	408
1262	QY	CGAGGTGAGGGACAGTGGGTGGCTGTGGAAATGTCCAGGTCCA-----CAGTAT	1212

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 14:56:31 ; Search time 296.5 Seconds
(without alignments)
5534.464 Million cell updates/sec

Title: US-09-975-253-1

Perfect score: 2699
Sequence: 1 ggttcagctgccgcagc.....aaaaaaaaaaaaaaaaaaaaa 1426

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q/cgn2_1/USPTO spoop/p/US09975253/runat 28102004 111509 19624/app query.fasta_1.1607
-DB=UniProt_02 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEX=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOCALIGN=200 -THR SCORE=opt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=score -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09975253@cgn 1 1 295@runat 28102004 111509 19624 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 02.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2458.5	91.1	452	Q7Z5G6	Q7Z5G6 homo sapien
2	2455.5	91.0	452	Q96GL3	Q96GL3 homo sapien
3	2326	86.2	427	1 IRF3 HUMAN	Q14653 homo sapien
4	2326	86.2	427	AAH71721	AAH71721 homo sapi
5	1742	64.5	419	2 Q764M6	Q764M6 sus scrofa
6	1742	64.5	419	2 BAD06317	BAD06317 sus scrofa
7	1553	57.5	419	1 IRF3 MOUSE	P70671 mus musculus
8	630	23.3	466	2 Q7ZXG6	Q7ZXG6 xenopus lae
9	504.5	18.7	488	2 Q6RCM8	Q6RCM8 homo sapien
10	504.5	18.7	488	2 AAR90325	Aar90325 homo sapi
11	504.5	18.7	488	2 AAR90326	Aar90326 homo sapi
12	503.5	18.7	497	1 IRF5 MOUSE	P56477 mus musculus
13	493.5	18.5	498	1 IRF5 HUMAN	Q13568 homo sapien
14	460.5	17.1	492	2 Q6PGZ7	Q6PGZ7 brachydanio
15	460.5	17.1	492	2 AAH56772	AAH56772 brachydan
16	460	17.0	460	2 Q6GR23	Q6GR23 xenopus lae

17	459.5	17.0	491	2	Q90ZD4	Q90ZD4 gallus gall
18	457.5	17.0	491	1	IRF3 CHICK	Q90ZD4 gallus gall
19	451.5	16.7	467	1	IRF6 HUMAN	Q14996 homo sapien
20	449.5	16.7	467	1	Q8WNO4	Q8WNO4 sus scrofa
21	448	16.6	503	1	IRF7 HUMAN	Q92985 homo sapien
22	447	16.6	460	2	Q6DDZ6	Q6DDZ6 xenopus lae
23	439	16.3	467	1	IRF6 MOUSE	P97431 mus musculus
24	438.5	16.2	459	2	Q57578	O57578 xenopus lae
25	437.5	16.2	467	2	Q9N136	Q9N136 ovis aries
26	436	16.2	457	1	IRF7 MOUSE	P70434 mus musculus
27	436	16.2	457	1	BAC37723	BAC37723 mus musculus
28	431	16.0	467	2	Q91VD0	Q91VD0 mus musculus
29	400.5	14.8	421	2	Q80ZH9	Q80ZH9 carassius a
30	400	14.8	392	2	Q6NY63	Q6NY63 brachydanio
31	400	14.8	392	1	AAH66724	AAH66724 brachydan
32	390.5	14.5	451	1	IRF4 HUMAN	Q15306 homo sapien
33	389.5	14.4	449	2	BAC40840	BAC40840 mus muscu
34	388	14.4	450	1	IRF4 MOUSE	Q64287 mus musculus
35	388	14.4	450	2	BAC40969	BAC40969 mus muscu
36	383.5	14.2	409	2	Q98TX6	Q98TX6 gallus gall
37	378.5	14.0	440	2	Q99419	Q99419 homo sapien
38	368.5	13.7	425	1	ICSB CHICK	Q90871 gallus gall
39	366.5	13.6	445	2	Q98TX7	Q98TX7 gallus gall
40	365.5	13.5	83	2	O54726	O54726 mus musculus
41	361.5	13.4	426	1	ICSB HUMAN	Q02556 homo sapien
42	359	13.3	423	2	Q6PE40	Q6PE40 brachydanio
43	359	13.3	423	2	AAH58298	AAH58298 brachydan
44	349.5	12.9	416	2	Q90W10	Q90W10 gallus gall
45	343.5	12.7	424	1	ICSB MOUSE	P23611 mus musculus

ALIGNMENTS

RESULT 1
Q7Z5G6 PRELIMINARY; PRT; 452 AA.
AC Q7Z5G6
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE IRF3 protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=23386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsu L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughlin N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay I.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gilts R.A.,
FAhey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;

RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC000660; AA00660.1; -	
DR	GO; GO:0005634; C:nucleus; IEA.	
DR	GO; GO:0003700; F:transcription factor activity; IEA.	
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001346; IRF.	
DR	Pfam; PF00605; IRF; 1.	
DR	ProDom; PD002355; IRF; 1.	
DR	SMART; SM00348; IRF; 1.	
DR	PROSITE; PS00601; IRF; 1.	
SQ	SEQUENCE 452 AA; 49134 MW; 96B059A029751B64 CRC64;	
Alignment Scores:		
Pred. No.:	1.95e-132	Length: 452
Score:	2458.50	Matches: 445
Percent Similarity:	98.45%	Conservative: 0
Best Local Similarity:	98.45%	Mismatches: 2
Query Match:	91.09%	Indels: 6
DB:	2	Gaps: 1
US-09-975-253-1 (1-1426) x Q7Z5G6 (1-452)		
QY	47 ATGGGAACCCCAAGACCCAGGNTCCCTCCCTGGCTGTCGCGAGCTGGACCTGGGGCAA	106
DB	1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGln	20
QY	107 CTGGAGGGGTGGCTGGGTGAACAAGAGCCGACCGCTTCGCGATCCCTTTGGAGACAC	166
DB	21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis	40
QY	167 GGCCTACGGCAGATGCACAGCAGGAGATTCGGATCTTCAGGCTGGGCGGAGGCC	226
DB	41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyLePheGlnAlaTrpAlaGluAla	60
QY	227 ACTGGTGATATGTTCCCGGAGGGGATAACGACAGCTGCCAACCTCGGAAGAGAAATTC	286
DB	61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe	80
QY	287 CGCTCTCCCTCAACCCCAAGAGGGTTCGTTTACGAGGACCGGACCAAGGACCTT	346
DB	81 ArgSerAlaLeuAsnArgLysGlyLeuArgLeuAlaGluAspArgSerLysAspPro	100
QY	347 CACGACCCACATAAATCTAGAGTTGTGAATCAGGAGTTGGGACTTTTCCAGCCA	406
DB	101 HisAspProHisLysIleTyTrpGluPheValAsnSerGlyValGlyAspPheSerGlnPro	120
QY	407 GACACCTCTCCGGACACCAATGGTGGAGGAGTCTCTGATACCCAGGAGACATCTG	466
DB	121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu	140
QY	467 GATGAGTTACTGGTAAACATGGTGTGGCCCCACTCCAGATCCGGGACCCCGAAGCTG	526
DB	141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu	160
QY	527 GCTTAGCCCTGAGCCCTGCCCTCAGCCCTGGGAGCCCGACTGGCAATCCCACT	586
DB	161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr	180
QY	587 CCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAAGCGGCTGTGGTCCCGGGGAA	646
DB	181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu	200
QY	647 GAGTGGGAGTTCGAGGTACAGCCTTCTACCGGGCGCCCAAGTCTCCAGCAGACCTC	706
DB	201 GluTrpGluPheGluValThrAlaPheTyArgGlyArgGlnValPheGlnThrIle	220
QY	707 TCCTCCCGGAGGCTCGGCTGGTGGTCCGAGTGGGAGGAGCAGCAGCTGCTGGA	766
DB	221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly	240
QY	767 TGGCCAGTCACACTGCCAGACCTGGCATGTCCCTGACAGACAGGAGGAGTGATGACTAC	826
DB	241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr	260
QY	827 GTGAGGCATGTGCTGAGCTGCTGGTGGGGGAACTGGCTCTCTGCGGGCGCCGCGAGTGG	886
DB	261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp	280
QY	887 CTCTGGGCCCCAGCGCTGGGGCACTGCCACATACATCTGGGAGTGGAGCAGTGTCTC	946
DB	281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyTrpAlaValSerGlyGluLeuLeu	300
QY	947 CCCAACAGCGGGCATGGCCCTGATGGCCAGTCCCCCAAGGCAAGAGAGGCGCTGT	1006
DB	301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGlyGlyValPhe	320
QY	1007 GACCTGGGGCCCTTCAATGT-----AGATCTGATTACCTTCACCGAAGG	1050
DB	321 AspLeuGlyProPheIleValGlySerTrpAlaProArgSerAspTyTrpHisGlyArg	340
QY	1051 AAGCGGACGCTCACACGCTATGCTCTGGTCTGTGGGGAGTCAAGCGCCAGGA	1110
DB	341 LysArgThrLeuThrThrLeuCysProLeuValLeuCysGlyGlyValMetAlaProGly	360
QY	1111 CCAGCCGTGGACCAAGAGCTGCTGATGTCAGGTTGTGCCACGCTGCTCAGGCGCTT	1170
DB	361 ProAlaValAspGlnGluAlaArgAspGlyGlnGlyCysAlaHisValProGlnGlyLeu	380
QY	1171 GTTAGAATGGCCCGGTAGGGGGTCCCTCTCTCCCTGGAGAATACTGTGGA-CTGCACAT	1230
DB	381 GlyArgAsnGlyProGlyArgGlyCysLeuLeuProGlyGluTyTrpCysGlyPr- AlaHis	400
QY	1231 TTCCAACAGCCACCCACTCTCCCTCACCTCCGACGACGACAGGCTTACTCTGAGGACTT	1290
DB	401 PheGlnGlnProProThrLeuProHisLeuArgProValGlnGlyLeuProAlaGlyLeu	420
QY	1291 GGTGGAGGCATGGAATTCAGGCGCTGGGAGAGCTGAGCCCTCGCTCTCATGTGT	1350
DB	421 GlyGlyGlyHisGlyPheProGlyProTrpGlyGluLeuSerProArgSerTrpCys	440
QY	1351 GCCTCCAAACCCCTCTTCCCCACACCTCAACCAA 1386	
DB	441 AlaSerAsnProProValProHisLeuAsnGln 452	
RESULT 2		
Q96GL3	PRELIMINARY; PRT; 452 AA.	
ID	Q96GL3	
AC	Q96GL3; 2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	IRF3 protein.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N. A.	
RC	TISSUE=Brain;	
RA	Medline=22388257; PubMed=12477932;	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.I., Scheetz T.E.,	
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahy J., Heltan E., Kettner M., Madan A., Rodriguez S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,	

RA Jones S.J., Marza M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RA Tissue=Brain;
 RL Strausberg R.;
 DR EMBL; BC009395; AA009395.1; -;
 DR HSSP; P23906; 21RF.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001346; 1RF
 DR InterPro; IPR008984; SMAD_FHA.
 DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam; PF00605; 1RF; 1.
 DR PRINTS; PD0267; INTFRNREGFT.
 DR ProDom; PD002355; 1RF; 1.
 DR SMART; SM00348; 1RF; 1.
 DR PROSITE; PS00601; 1RF; 1.
 SQ SEQUENCE 452 AA; 49120 MW; 96B059A028AE6B64 CRC64;

Alignment Scores:
 Pred. No.: 2,896-132 Length: 452
 Score: 2455.50 Matches: 444
 Percent Similarity: 98.45% Conservative: 1
 Best Local Similarity: 98.23% Mismatches: 2
 Query Match: 90.98% Indels: 6
 DB: 2 Gaps: 1

US-09-975-253-1 (1-1426) x Q96GL3 (1-452)
 QY 47 ATGGGAAACCCCAAGCCAGGAGGATTCCTGCTGGTGTGCGAGCTGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgLeuLeuProTyrLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGGAGGGCGTGGCCCTGGTGTACAGAGCCGCGCTTCGGATCCCTTGGAGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgLeuProTyrLysHis 40
 QY 167 GGCCTACGCGAGGATGACAGCAGGAGGATTCGGAACTTCAGGCGCTGGCGCCAGGCG 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyLeuPheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTGCATATGTTCCCGGGAGGATAGCCAGACCTGCCAACCTGGAGAGGAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCGCTCAACCGCAAGAGGTTGCGTTTACGAGAGCGGAGCAGGACCCCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAATCTACAGTTTGTGAATCTCAGGAGTTGGGACTTTTCCAGGCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCCGGACACCAATGGTGGAGCAGTACTTCTGATACCCAGGAGACATTCCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspLeu 140
 QY 467 GATGAGTTACTGGGTAACTGTGTGGCCCACTCCAGATCCCGGAGCCCAATCCCACT 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCGCTCAGCCCTGCGGAGCCCGAGCTTCGACAAATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGAGCGGCTGTGGTGGCGGGGAA 646
 Db 181 PropheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200

QY 647 GAGTGGAGTTCGAGGTGACAGCCTTCTACCGGGCGCCCAAGTCTTCCAGACACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCCGAGGCGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTTCACACTGCGCAGACCTGCGTGTCTCCCTGACACAGGAGGAGTGTAGAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGCATGTGTGAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGGCCGCGCTGGGGCTGCTCCACATCTGCGCAGTGTAGCGAGGAGTGTCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeu 300
 QY 947 CCAACAGCGGGCATGCGGCTGATGCGGAGTCCCAAGGACCAAGGAGGAGGCGTGT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGCGCTTCAATTGT-----AGATCTGATTACCTTACCGGAAGG 1050
 Db 321 AspLeuGlyProPheIleValGlySerTrpAlaProArgSerAspTyrLeuHisGlyArg 340
 QY 1051 AAGCGAGCCTCACACGCTATGCCCTCTGTTCTGTGGGGAGTGTATGGCCCGGAGG 1110
 Db 341 LysArgThrLeuThrThrLeuCysProLeuValLeuCysGlyGlyValMetAlaProGly 360
 QY 1111 CCAGCGCTGAGCAACAGAGGCTGCTGATGCTCAAGTTGTGCGCCACGCTGCTCAGGCGCTT 1170
 Db 361 ProAlaValAspGlnGlnAlaArgAspGlyGlnGlyCysAlaHisValProGlnGlyLeu 380
 QY 1171 GGTAGAAATGGCGCGGTAGGGGTGCTCTCTCCCTGGAGAAATACTGTGGACCTGCACAT 1230
 Db 381 GlyArgAsnGlyProGlyArgGlyCysLeuLeuProGlyGlyTyrCysGlyProAlaHis 400
 QY 1231 TTCCACACGACCCACTCTCCCTCAGTCCGACGACGATCAAGGCTTACCTGAGGACTT 1290
 Db 401 PheGlnGlnProProThrLeuProHisLeuArgProValGlnGlyLeuProAlaGlyLeu 420
 QY 1291 GGTGAGGCGCATGGATTTCAGGGCGCTGGGAGAGCTGAGCCCTGCTCCTCCTGCTGT 1350
 Db 421 GlyGlyGlyHisGlyPheProGlyProThrGlyAspLeuSerProArgSerTrpCys 440
 QY 1351 GCCTCCAAACCCCTGTTTCCCAACCCACCTCAACCAA 1386
 Db 441 AlaSerAsnProProValProHisLeuAsnGln 452

RESULT 3
 IRP3_HUMAN
 ID IRP3_HUMAN STANDARD; PRT; 427 AA.
 AC Q14653;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Interferon regulatory factor 3 (IRF-3).
 GN Name=IRF3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA MEDLINE=96102173; PubMed=8524823;
 RX Au W.-C., Moore P.P.A., Lowther W.W., Juang Y.-T., Pitha P.M.;
 RT "Identification of a member of the interferon regulatory factor family
 RT that binds to the interferon-stimulated response element and activates


```

Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
QY 1307 TTCAGGGCCCTGGGAGAC 1327
Db 421 PheGlnGlyProGlyGluSer 427

RESULT 4
AAH71721
ID AAH71721 PRELIMINARY; PRT; 427 AA.
AC AAH71721.
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE IRF3 protein.
OS IRF3.
GN IRF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Toshiyuki S., Carninci P., Prange C.,
RA Bosak S.A., McEwan P.J., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahy J., Helton E., Sodergren E.J., Lu X., Gibbs R.A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071721; AAH71721.1; -.
SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;

Alignment Scores:
Pred. No.: 7,21e-125 Length: 427
Score: 2326.00 Matches: 426
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 86.18% Indels: 0
DB: 2 Gaps: 0

US-09-975-253-1 (1-1426) x AAH71721 (1-427)

QY 47 ATGGGAACCCCAAGCCACGGNTCTCGCTGGCTGGTGTGGAGCTGGACCTGGGGCAA 106
Db 1 MetGlyThrProLysProArgGlyLeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
QY 107 CTGAGGGCGTGGCTGGGTGAACAAGACGGCCGCGCTCCGCATCCCTCGATCGAAGCAC 166
Db 21 LeuGluGlyValAlaTpvAlaAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
QY 167 GGCTACGGCAGGTGACACAGCAGGAGATTTCGAATCTTCCAGGCTGGCGCAGGCC 226
Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTpvAlaGluAla 60

```

```

QY 227 ACTGTGTCATATGTTCCCGGAGGGATAAGCCAGACCTGCCAACCTGCAAGAGAAATTC 286
Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAspPhe 80
QY 287 CGCTCTGCCCTCAACCGCAAGAGGTTCGTTTAGCAGAGACCGGACGAGGACCCCT 346
Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY 347 CACGACCCACATAAAATCTACGAGTTTGTGAATCTCAGGAGTTGGGACTTTTCCAGCCA 406
Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY 407 GACACCTCTCCGGACACCAATGGTGAGGACGAGTACTTCTGATACCCAGGAAGACATCTG 466
Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
QY 467 GATGAGTTACTGGTAAATGGTGTGGCCCACTCCAGATCCCGGAGCCCAAGCCTG 526
Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGCCGAGCCCGAGTGGACATCCCACT 586
Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY 587 CCTTCCCAAACTCGGGCCCTCTGAGAACCCACTGAGCGGCTGTGTGTCGGGGGAA 646
Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
QY 647 GAGTGGGAGTTCGAGGTGACAGCCTTACCGGGCCGCCCAAGTCTTCAGCAGACCAATC 706
Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnThrIle 220
QY 707 TCCTCCCGGAGGGCTCGGCTCGTGGTCCGAGTCCGAGTGGGAGACAGACGCTGCTGA 766
Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY 767 TGGCCAGTCACACTGCCAGACCTGGCATGTCCCTGACAGACAGGGGAGTGTAGCTAC 826
Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
QY 827 GTGAGGCATGTCTGAGTGTGGTGGGAGCTGGCTCTCTGCGCGCCGCGGAGTGG 886
Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY 887 CTCTGGGCCCCAGCGCTGGGCACTGCCACATCTGGGAGTGGGAGTGGGAGGAGTGTCTC 946
Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluLeuLeu 300
QY 947 CCCAACAGCGGGCATGGGCTGTGATGGGAGGTCCCAAGGACAAAGGAGGAGGCTGTTT 1006
Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
QY 1007 GACTGGGGCCCTTCATGTAGATCTGATTACCTTCCAGGAAGGAGCGGACGCTCACA 1066
Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
QY 1067 CGCTATGCCCTCTGTTCTGTGGGGAGTTCATGGCCCGGAGGACGAGCCGTGGACCAAG 1126
Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
QY 1127 AGGCTCGTGTGTTCAAGGTGTGGCCACCTGCTCAGGGCTTGTGTAGAAATGCGCCGG 1186
Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
QY 1187 GTAGGGGTGCTCTCTCCCTGGAGAAATCTGTGGACCTGCACATTTCCACAGCCACCCA 1246
Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
QY 1247 CTCTCCCTCACTCCAGCAGTACAGGCTACTCGAGGACTTGTGTGGAGGCGCATGGAT 1306
Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420

```

Fri Oct 29 11:11:45 2004

QY	1307	TTCCAGGGCCCTGGGAGAGC	1327	
Db	421	PheGlnGlyProGlyGluSer	427	
RESULT 5				
Q764M6		PRELIMINARY;	PRT;	419 AA.
ID	Q764M6			
AC	Q764M6			
DT	05-JUL-2004	(TRENBLrel. 27, Created)		
DT	05-JUL-2004	(TRENBLrel. 27, Last sequence update)		
DT	05-JUL-2004	(TRENBLrel. 27, Last annotation update)		
DE		Interferon regulatory factor 3.		
GN	Name=irf3;			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
NCBI_TaxID=9823;				
RN	[1]			
PP	SEQUENCE FROM N.A.			
RA	Uenishi H., Eruchi T., Suzuki K., Sawazaki T., Toki D., Shinkai H.,			
RA	Okumura N., Hamashina N., Awata T.,			
RT	"PEDE (Pig EST Data Explorer): construction of a database for ESTs			
RT	derived from porcine full-length cDNA libraries."			
RL	Nucleic Acids Res. 32:D484-D488(2004).			
DR	EMBL; AB116563; BAD06317.1; ..			
DR	InterPro; IPR001346; IRF.			
DR	InterPro; IPR008984; SMAD FHA.			
DR	InterPro; IPR009058; Wing_hlx_DNA_bnd.			
DR	Pfam; PF00605; IRF; 1.			
DR	PRINTS; PRO0267; INTERFERGFC.			
DR	ProDom; PD002355; IRF; 1.			
DR	SMART; SM00348; IRF; 1.			
DR	PROSITE; PS00601; IRF; UNKNOWN 1.			
SQ	SEQUENCE 419 AA; 46645 MW; 8B731861BE002284 CRC64;			
Alignment Scores:				
Pred. No.:	1.68e-91	Length:	419	
Score:	1742.00	Matches:	327	
Percent Similarity:	83.61%	Conservative:	25	
Best Local Similarity:	77.67%	Mismatches:	67	
Query Match:	64.54%	Indels:	2	
DB:	2	Gaps:	1	
US-09-975-253-1 (1-1426) x Q764M6 (1-419)				
QY	47	ATGGGAACCCCAAGCCAGGTCCTCCCTGGCTGGTGTGGAGCTGGAGCTGGGCAA	106	
Db	1	MetGlyThrGlnLysProArgGlyLeuProTrpLeuLeuSerGlnLeuAsnGlnGlyGln	20	
QY	107	CTGAGGGCGTGGCTGGTGAACAAGAGCGCAGCGCTTCCGCATCCCTTGAAGCAC	166	
Db	21	LeuGluGlyValAlaTrpLeuAspGluGlyHisThrArgPheArgGlyProTrpLysHis	40	
QY	167	GGCTACGGCAGATGACACAGAGGATTCGGATCTCCAGGCTGGGCGGAGGCC	226	
Db	41	GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyLeuPheGlnAlaTrpAlaGluAla	60	
QY	227	ACTGGTGCATATGTTCCCGGAGGATTAAGCCAGACCTGCCAACCTGGAAGAGGAATTC	286	
Db	61	SerGlyAlaIleThrProGlyLysAspLysProAspLeuProThrTrpLysArgAsn	80	
QY	287	CGCTCGCCCTCAACCGCAAGAGGGTGGCTTTAGAGAGGACCGGAGGACCCCT	346	
Db	81	ArgSerAlaLeuAsnArgLysGluAlaLeuArgLeuAlaGluAspHisSerLysAsp	100	
QY	347	CACGACCCACATAAATCTACAGTTGTGAACTCAGAGTTGGGACTTTTCCAGGCA	406	
Db	101	HisAspProHisLysIleThrGluPheValThrSerGlyValGlyAspPheProGlu	120	
QY	407	GACACTCTCCGAGACCAACCTGGTGGAGGAGTCTTCTGTATACCCAGAGACATTC	466	
Db	121	AspThrSerLeuAspLeuSerGlyArgTyrSerThrSerAspThrHisGluAspSer	140	
QY	467	GATGAGTTACTGGTAACTAGTGTGGTGGCCCTCCAGATCCGGGACCCCAAGCCTG	526	
Db	141	AspLysLeuLeuSerGlyMetAspLeuA-----SerAspAlaGlyProGlnSerLeu	158	
QY	527	GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTCGGAGCCCTCAGCTTGGACAATCCCACT	586	
Db	159	ThrLeuAlaLeuGlnProGlnLeuSerLeuSerProSerValAspAlaProAla	178	
QY	587	CCCTTCCCAACCTGGGCGCCCTCTGAGAACCCTGAAAGCGGCTGTGGTCCCGGGGAA	646	
Db	179	SerCysProLeuLeuGlyValArgGluAsnProLeuLysGlnLeuLeuAlaAsnAsp	198	
QY	647	GAGTGGAGTTTCAGGTGACAGCCCTTACCGGGCGCCCAAGTCTCCACAGACCAATC	706	
Db	199	GluTrpGluPheGlnValThrValPheTyrArgGlyCysGlnValPheGlnGlnThr	218	
QY	707	TCCTGCCCGGAGGCGCTGGCTGGTGGTCCGAAGTGGGAGAGCAGACGCTGCTGGA	766	
Db	219	CysSerProGlyGlyLeuArgLeuValGlySerGluAlaGluAspGlyThrLeuAla	238	
QY	767	TGGCCAGTCACACTGCCAGACCTGGCATGTCCTGTACAGACAGGGGAGTGTAGCTAC	826	
Db	239	GlnProValArgLeuProAspProAlaAlaSerLeuThrAspArgGlyValAlaAsp	258	
QY	827	GTGAGGCATGCTGAGCTGCTGGTGGGAGTGGCTCTCTGCGGGCGCGGAGTGG	886	
Db	259	ValArgArgValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp	278	
QY	887	CTCTGGGCGCGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGC	946	
Db	279	LeuTrpAlaGlnArgLeuGlyHisCysHisValTyrTrpAlaMetGlyGluGluLeu	298	
QY	947	CCCAACAGCGGCGCTGGGCTGATGGCGAGTCCCAAGTCCCAAGCAAGAGGAGCGCT	1006	
Db	299	ProAspSerGlyHisLysProAspGlyGluValProLysAspArgGluGlyValPhe	318	
QY	1007	GACCTGGGCGCTTCTATTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG	1066	
Db	319	AspLeuGlyProPheIleGluAspLeuAlaPheIleGluIleSerArgArgSerPro	338	
QY	1067	CGCTATCCCTCTGGTCTGTGGGGAGTATGTCGCGGAGTATGTCGCGGAGTATG	1126	
Db	339	ArgTyrThrLeuTrpPheCysMetGlyGlnSerTrpProGlnAspGluProTrpVal	358	
QY	1127	AGGCTCGTATGCTGAAGTGTGTGCCACCTGCTGAGGCGCTGTGTAAGTGGCCCG	1186	
Db	359	ArgLeuValMetValLysValValProMetCysLeuArgAlaLeuValAspMetAla	378	
QY	1187	GTAAGGGGTGCTCTCCCTGGAGATATCTGTGGACCTGCACATTTTCCACAGCCCA	1246	
Db	379	AspGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisLysSerAsnSerHis	398	
QY	1247	CTCTCCCTCACCTCCGACCACTGACAGGCTTACCTGACAGGCTTGTGGAGGGCAT	1306	
Db	399	LeuSerLeuThrSerAspGlnTyrLysAlaCysLeuArgAspLeuValGluAspMet	418	
QY	1307	TTC 1309		
Db	419	Phe 419		
RESULT 6				
BAD06317		PRELIMINARY;	PRT;	419 AA.
ID	BAD06317			
AC	BAD06317			
DT	02-MAR-2004	(TRENBLrel. 27, Created)		
DT	02-MAR-2004	(TRENBLrel. 27, Last sequence update)		
DT	02-MAR-2004	(TRENBLrel. 27, Last annotation update)		
DE		Interferon regulatory factor 3.		
GN	IRF3.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
NCBI_TaxID=9823;				

```

[1]
RP SEQUENCE FROM N.A.
RA Uneshi H., Eguchi T., Suzuki K., Sawazaki T., Toki D., Shinkai H.,
RA Okumura N., Hamashima N., Awata T.:
RT "PEDE (Pig EST Data Explorer): construction of a database for ESTs
RT derived from porcine full-length cDNA libraries.";
RL Nucleic Acids Res. 32:D484-D488 (2004).
DR EMBL; AB116563; BAD06317.1; -.
SQ SEQUENCE 419 AA; 46645 MW; 8B731861BE002284 CRC64;

Alignment Scores:
Pred. No.: 1 68e-91 Length: 419
Score: 1742.00 Matches: 327
Percent Similarity: 83.61% Conservative: 25
Best Local Similarity: 77.67% Mismatches: 67
Query Match: 64.54% Indels: 2
DB: 2 Gaps: 1

US-09-975-253-1 (1-1426) x BAD06317 (1-419)
QY 47 ATGGAAACCCAAAGCCACGGNCTCGCCTGGCTGGTGTGCAGCTGACCTGGGGCAA 106
Db 1 MetGlyThrGlnLysProArgIleLeuProTrpLeuIleSerGlnLeuAnGlnGlyGln 20
QY 107 CTGGAGGGCGTGGCTGGTGAAACAAGAGCCGACCGGCTCCGATCCCTTTGGAGCAC 166
Db 21 LeuGluGlyValAlaTrpLeuAspGluGlyHisThrArgPheArgIleProTrpLysHis 40
QY 167 GGCCTACGGCAGGATGCACAGCAGGAGGATTTCCGAATCTTCCAGGCTTGGGCCAGGCC 226
Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
QY 227 ACTGGTGCATATGTTCCCGGAGGAGTAAAGCAGACCTGCCAACTCGAAGAGGAAATTC 286
Db 61 SerGlyAlaTrpThrProGlyLysAspLysProAspLeuProThrTrpLysArgAnPhe 80
QY 287 CGCTCTGCCCTCAACCGCAAGAGGGTTGGTTAGCAGAGCACCGAGCAAGACCCCT 346
Db 81 ArgSerAlaLeuAsnArgLysGluAlaLeuArgLeuAlaGluAspHisSerLysAspPro 100
QY 347 CACCACCCACATAAAATCTACGAGTTGTGAACCTCAGAGTTGGGACATTTTCCAGCA 406
Db 101 HisAspProHisLysIleTyrGluPheValThrSerGlyValGlyAspPheProGluPro 120
QY 407 GACACCTCTCGGACACCAATGGTGGAGGCACTCTCTGATACCCAGCAAGACATTCGT 466
Db 121 AspThrSerLeuAspLeuSerGlyArgTyrSerThrSerAspThrHisGluAspSerLeu 140
QY 467 GATGAGTACTGGGTAAATGGTGTGGCCCCCATCTCCAGATCCGGGACCCCAAGCCTG 526
Db 141 AspLysLeuLeuSerGlyMetAspLeuAla-----SerAspAlaGlyProGlnSerLeu 158
QY 527 GCTGTAGCCCTGAGCCCTGCGCTCAGCCCTCGCGGAGCCCGACTGGACAATCCCCT 586
Db 159 ThrLeuAlaLeuGluGlnProProGlnLeuSerLeuSerProSerValAspAlaProAla 178
QY 587 CCCTTCCAAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGCTGTGGTCCCGGGGAA 646
Db 179 SerCysProAsnLeuGlyValArgGluAsnProLeuLysGlnLeuLeuAlaAsnAspAsp 198
QY 647 GAGTGGGAGTTCAGAGTGACAGCCTTCTACGGGGCCGCCAAGTCTTCCAGCAGACCATC 706
Db 199 GluTrpGluPheGlnValThrValPheTyrArgGlyCysGlnValPheGlnGlnThrVal 218
QY 707 TCCTGCGCGGAGGGCGCTGGCTGGTGGGGTCCGAAGTGGGAGACAGACGCTGCCTGGA 766
Db 219 CysSerProGlyGlyLeuArgLeuValGlySerGluAlaGluAspGlyThrLeuAlaGly 238
QY 767 TGGCCAGTACACTGCCAGACCCCTGGCATGCTCCCTGACAGACAGGGGAGTGTAGCTAC 826
Db 239 GlnProValArgLeuProAspProAlaAlaSerLeuThrAspArgGlyValAlaAspTyr 258
QY 827 GTGAGGCATGTGCTAGCTCCCTGGGTGGGGGAGCTGGCTCTCTGGCGGGCGGACGTGG 886

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences".
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: May activate transcription by complex formation with
 CC other transcriptional factors, possibly members of the STAT
 CC family. Binds specifically to the IRF-1 stimulated response element
 CC (ISRE) but not to the IRF-1 binding site PRD-1 (by similarity).
 CC -2- SUBCELLULAR LOCATION: Nuclear.
 CC -3- SIMILARITY: Belongs to the IRF family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

Alignment Scores:
 Pred. No.: 1,06e-80 Length: 419
 Score: 1553.00 Matches: 301
 Percent Similarity: 78.59% Conservatives: 33
 Best Local Similarity: 70.82% Mismatches: 83
 Query Match: 57.54% Indels: 8
 DB: 1 Gaps: 6

us-09-975-253-1 (1-1426) x IRF3_MOUSE (1-419)

QY	47	ATGGGAACCCCAAGCCACGAGTCTCCCTCGCTGGTGTGGAGCTGGACCTGGGGCAA	106
Db	1	MetGluThrProLysProArgGluLeuProTrioLeuValSerGlnLeuAspLeuGlyGln	20
QY	107	CTGGAGGCGTGGCTGGTGAACAGAGCGCGCGCTTCCGCATCCCTTGGAGCAC	166
Db	21	LeuGluGlyValAlaTrioLeuAspGluSerArgThrArgPheArgGluProTrioLysHis	40
QY	167	GGCTACGGCAGGATGACAGCAGGAGATTTCCGAATCTTCCAGGCTGGCGCCAGGCC	226
Db	41	GlyLeuArgGlnAspAlaGlnMeAlaAspPheGlyllePheGlnAlaTrioAlaGluAla	60
QY	227	ACTGCTGCATATGTTCCCGGAGGAGTAAGCCAGACCTGCCAACCTGGAAGAGGAATTC	286
Db	61	SerGlyAlaTrioThrProGlyLysAspLysProAspValSerThrTrioLysArgAsnPhe	80
QY	287	CGCTCTGCTCCCAACCGCAAGAGGGTTGGTTTGTAGCAGAGGACCGGACGACGACCT	346
Db	81	ArgSerAlaLeuAsnArgLysGluValLeuArgLeuAlaAlaAsnSerLysAspPro	100
QY	347	CACGACCCACATAAATCTACGAGTTTGTGAATCTCAGGAGTTGGGAGCTTTTCCAGCCA	406
Db	101	TyrAspProHisLysValTrioLysValThrProGlyAlaArgAspPheValHisLeu	120
QY	407	GACACCTCTCCCGACACCAATGGTGGAGCGACTCTCTGTATACCGAGCAACATTCTG	466
Db	121	GlyAlaSerProAspThrAsnGlyLysSerSerLeuProHisSerGlnGluAsn---Leu	139

QY	467	GATGAGTTACTGGGTAACTAGTGTGGCCCTCCAGATCCGGAGACCCCAAGCGCTG	526
Db	140	ProLysLeuPheAspGlyLeuLeuLeuGlyProLeuLysAspGluGlySerSerAspLeu	159
QY	527	GCTGTAGCCCTGAGCCCTCCCTCAGCCCTCGGAGCCCTCAGCTTGGACAAATCCCACT	586
Db	160	AlaIleValSerAspPro---SerGlnGlnLeuProSerProAsnValAsnAsn----	176
QY	587	CCCTTCCCAACCTGGCGCCCTCTGAGAACCCACTGAAGCGGCTGTGGTCCCGGGGAA	646
Db	177	---PheLeuAsnProAlaProGlnGluAsnProLeuLysGlnLeuLeuAlaGlu---Glu	194
QY	647	GAGTGGAGTTCGAGGTGACAGCCCTTACCGGGCGCCCAAGTCTCCAGCAGACCATC	706
Db	195	GlnTrioGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrLeu	214
QY	707	TCCTGCCCGGAGGCGCTGGCGCTGGTGGGTCCGAAGTGGGAGACAGCAGCTGCTGGA	766
Db	215	PheCysProGlyGlyLeuArgLeuValGlySer---ThrAlaAspMetThrLeuProTrio	233
QY	767	TGGCCAGTCCACTGCCAGACCTGGCATGTCCCTGACAGACAGGGGAGTGTAGCTAC	826
Db	234	GlnProValThrLeuProAspProGluGlyPheLeuThrAspLysLeuValLysGluTyr	253
QY	827	GTGAGGCATGTGCTGAGTGTGGTGGGAGCTGGCTCTCTGGCGGCGCGGAGTGG	886
Db	254	ValGlyGlnValLeuLysGlyLeuGlyAsnGlyLeuAlaLeuTrioPheGlnAlaGlyGlnCys	273
QY	887	CTCTGGCGCCAGCGCTGGCGCTGGCATGTGGAGTCCCAACATATCTGGCGAGTGGACGAGTGTCT	946
Db	274	LeuTrioPheGlnArgLeuGlyHisSerHisAlaPheTrioAlaLeuGlyGluLeuLeu	293
QY	947	CCCAACAGCGGCGCATGGCTGTGGTGGAGTCCCGAGTCCCAAGCAGACAGGAGGAGCGCTGTTT	1006
Db	294	ProAspSerGlyArgGlyProAspGlyGluValHisLysAspLysAspGlyAlaValPhe	313
QY	1007	GACCTGGCGCCCTTCAATGTAGATCTGATTACTCTTCCAGGAGGAGGAGCGCTCACCA	1066
Db	314	AspLeuArgProPheValAlaAspLeuIleAlaPheMetGluGlySerGlyHisSerPro	333
QY	1067	CGCTATGCTCTGCTGTGTGGGAGTCTATGCTGGCGGAGTCTATGCTGGCGGAGTGGAGTGG	1126
Db	334	ArgTyrThrLeuTrioPheCysMetGlyGluMetTrioProGlnAspGlnProTrioValLys	353
QY	1127	AGCTCTGCTGATGTGAAGTGTGGCCACGCTGCTCAGGGCCTTGTAGAAATGGCCCGG	1186
Db	354	ArgLeuValMetValLysValValProThrCysLeuLysGluLeuLeuGluMetAlaArg	373
QY	1187	GTAGGGGTGCTCTCTCTCTGGAGATCTGTGGACCTGCACATTTTCCAGACGACCCCA	1246
Db	374	GluGlyGlyAlaSerSerLeuLys---ThrValAspLeuHisLysSerAsnSerGlnPro	392
QY	1247	CTCTCCCTCAGCTCCGACCATCAAGGCTTACTCTGAGGACTTGTGGAGGCGATGGAT	1306
Db	393	IleSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluAspMetAsp	412
QY	1307	TTCAGGCGCTCTGGG 1321	
Db	413	PheGlnAlaThrGly 417	

RESULT 8
 Q7ZXG6 PRELIMINARY; PRT; 466 AA.
 ID Q7ZXG6;
 AC Q7ZXG6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MG53080 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;

35 TrpValAsnGlyGluIysLysLeuPheCysIleProTirparghiSLAAlaThrArgHisGly 54

182 GCACAGCAGGAG---GATTTCCGGNAATCTTCAGGCCTGGCGCAGGCCACTGTGGCATAT 238

55 ProSerGlnAspGlyAspAsnThrIlePhePhylsAlaIrpAlaIysGluThrGlyLysTyr 74

239 GTTCCCAGGAGGATAACCCAGACCTGCCAACCTCGAAGAGAAATTCCCGTCTCCCTC 298

75 ThrGluGlyValAspGluAlaAspProAlaIysTrpLysAlaAsnLeuArgCysAlaLeu 94

299 AACCGCAAGAAGGTTTGCTTTAGCAGAGCAGGACAGGAC---CCTCACGACCACA 355

95 AsnLysSerArgAspPheArgLeuIleTyrAspGlyProArgAspMetProGlnPro 114

356 CATAAATCTACGAGTTGTGAACCTCAGAGTTGGGGACTTTTCCACGACACACCTCT 415

115 TyrIlysIleTyrGluValCysSerAsnGlyPro-----AlaProThrAspSerGln 131

416 CCG-----GACACCAATCGTGGAGCGAGTAGTCTCTGATACCCAGGAAGACATTCTGGAT 469

132 ProProGluAspTyrSerPheGlyAlaGlyGluGluGluGluGluGluLeuGln 151

470 GAGTTACTGGTPAACATG-----GTGTTGGCCCCA 499

152 ArgMetLeuProSerLeuSerLeuThrGluAspValLysTrpProThrLeuGlnPro 171

500 CTCACAGATCCGGGACCCCAAGCCTGGCTGTAGCCCTCAGCCCTCGCCCTCAGCCCTG 559

172 ---ProThrLeuGlnProProValValLeuGlyProProAlaProAspProSerProLeu 190

560 CGAGGCCCGAGCTTGGACAAATCCACTCCCTTCCCAAACCTGGGGGCCCTCTGAG----- 613

191 ---AlaProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu 208

614 -----AACCCTGAAGCGCTGTGTGTCGCG---GGGCAAGAG----- 649

209 GluProGlyProLeuProAlaSerLeuProProAlaGlyGluGlnLeuLeuProAspLeu 228

650 -----TGGGAGTTCAGGTGACAGCTGACAGCCTTCTACCGGGGC 682

229 LeuIleSerProHisMetLeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGly 248

683 CGCCAAGTCTTCCAGCAGACCATCTCTCGCCGAGGGCGCTGGCTGGTGGGTCGCA 742

249 ArgProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrSerGln 268

743 GTG-----GGAGCAGGACGCTGCCTGGATGGCCA 772

269 LeuGluAlaThrGlnGluInValGluLeuPheGlyProIleSerLeuGlu-----Gln 286

773 GTCACATGCCAGACCTGGCATCTCCCTGACACAGCAGGAGGATGATGAGTACGTGAGG 832

287 ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrThrAsn 305

833 CATGTGTGAGCTGCCTGGTGGGGAGCTGCTCTCTGCGGGCGGGCAGTGGCTCTGG 892

306 GlnLeuLeuaspValLeuAspArgGlyLeuIleLeuGlnLeuGlnGlyGlnAspLeuTyr 325

893 GCCCAGCGGCTGGGGCAGCTGCCACATACTGG-----GCAGTGCAGCGGAGGAG 940

326 AlaIleArgLeuCysGlnCysLysValPheIleTrpSerGlyProCysAlaSerAlaHisAsp 345

941 CTGCTCCCAACAGCGGGCATGGGCTGATGCGAGGTCCTCCCAAGGACAAAGAGAGGC 1000

346 SerCysProAsn-----ProIleGlnArgGluValIysThrLys 358

1001 GTGTTTGACCTGGGGCCCTTATTGTAGATCTGATTACCTTCACGGAAGCA---AGCGGA 1057

359 LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn 378

1058 CGCTCACACCGTATGCCCTCTGGTTCTGTGGGGGAGTCTATGGCCCGGACGACCG 1117

379 ThrProProPheGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro 398

QY	1118	TTGACCAAGAGCTCGTGTAGTGTCAAGCTTGTGCCACCGTGCCTCAGGGCTTCGTAGAA	1177
Db	399	ArgGluLysLeuLeuIleThrValGlnValValProValAlaAlaArgLeuLeuLeuGlu	418
QY	1178	ATGGCCCGGTAGGGGTGCCTCTCCCTCGGAGAAATCTGGACCTGCACATTTCCAAC	1237
Db	419	MetPheSer--GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn	437
QY	1238	AGCCA-----CCCACTCTCCCTCAC-----	1257
Db	438	--ProAspLeuLysAspArgMetValGluGlnPheLysGluLeuHisIleTrpGlnS	457
QY	1258	-----CTCGACCACTACAAGCCCTACTCTGCAGGACTTGTGGAGGCGATCGAT	1306
Db	457	erGlnGlnArgLeuGlnProValAlaGlnAlaProProGlyAlaGlyLeuGlyValGly-	476
QY	1307	TTCCAGGGCCCTGG	1320
Db	477	--GlnGlyProTrp	480
RESULT 10			
AAR90325			
ID	AAR90325	PRELIMINARY;	PRT; 488 AA.
AC	AAR90325;		
DT	02-MAR-2004	(TRENBLrel. 27, Created)	
DT	02-MAR-2004	(TRENBLrel. 27, Last sequence update)	
DT	02-MAR-2004	(TRENBLrel. 27, Last annotation update)	
DE	Interferon regulatory factor 5 variant 3.		
GN	IRF5.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21316468; PubMed=11303025;		
RA	Barnes B.J., Moore P.A., Pitha P.M.;		
RT	"Virus-specific activation of a novel interferon regulatory factor,"		
RT	IRP-5, results in the induction of distinct interferon alpha genes.,"		
FL	J. Biol. Chem. 276:23382-23390(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Barnes B.J., Pitha P.M.;		
RL	Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AY504946; AAR90325.1; -		
SQ	SEQUENCE 488 AA; 54942 MW; 4E43B54EE7A7E37 CRC64;		
Alignment Scores:			
Pred. No.:	8,74e-21	Length:	488
Score:	504.50	Matches:	153
Percent Similarity:	44.47%	Conservative:	62
Best Local Similarity:	32.21%	Mismatches:	186
Query Match:	18.69%	Indels:	95
DB:	2	Gaps:	22
US-09-975-253-1 (1-1426) x AAR90325 (1-488)			
QY	2	GTTCACAGCTGCCCGACGCCCGACCTTCCATCTGAGCCGACCATGGGAACCCCAAG	61
Db	5	IleProValAlaProThrProArgArgVal-----	15
QY	62	CCACGGNTCTCGCTCGCTGTGTGCGAGCTGGACCTGGGGCACTGGAGGGCGTGGCC	121
Db	16	---ArgLeuLysProTrpLeuValAlaGlnValAsnSerCysGlnTyrProGlyLeuGln	34
QY	122	TGGGTGAACAAGACGGCAGCGCTTCGCGCATCCCTTGGAAACGACGGCTACGGCAGAT	181
Db	35	TrpValAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly	54
QY	182	GCACAGCAGGAG--GATTTCGATATCTTCAGGCTGGCGGCGGACCATGTGTGCATAT	238
Db	55	ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTr	74

Fri Oct 29 11:11:45 2004

```

QY 1178 ATGCGCGGTAGGGGTGCTCTCTCCCTGAGAACTACTGTGACCTGCACTTTTCAAC 1237
Db 1419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
QY 1238 AGCCA-----CCCACTCTCCCTCAC-----1257
Db 438 --ProAspLeuLeuAspArgMetValGluGlnPheLysGluLeuHisHisIleTrpGlnS 457
QY 1258 -----CTCGACCACTACAAGCCCTACCTGAGGAGCTTGGTGGAGGGCATGGAT 1306
Db 457 erGlnGlnArgLeuGlnProValAlaGlnAlaProGlyAlaGlyLeuGlyValGly- 476
QY 1307 TTCCAGGGCCCTGG 1320
Db 477 --GlnGlyProTrp 480

RESULT 11
AAR90326 PRELIMINARY; PRT; 488 AA.
AC AAR90326;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Interferon regulatory factor 5 variant 4.
GN IRF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21316468; PubMed=11303025;
RA Barnes B.J.; Moore P.A.; Pitha P.M.;
RT "Virus-specific activation of a novel interferon regulatory factor,"
RT IRF-5, results in the induction of distinct interferon alpha genes.";
RT J. Biol. Chem. 276:23382-23390 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Barnes B.J.; Pitha P.M.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504947; AAR90326.1; -.
SQ SEQUENCE 488 AA; 54942 MW; 4E43B54EE74A7B37 CRC64;

Alignment Scores:
Pred. No.: 8,74e-21 Length: 488
Score: 504.50 Matches: 163
Percent Similarity: 44.47% Conservative: 62
Best Local Similarity: 32.21% Mismatches: 186
Query Match: 18.69% Indels: 95
DB: 2 Gaps: 22

US-09-975-253-1 (1-1426) x AAR90326 (1-488)
QY 2 GTTCCAGCTGCGGACGCCCGGACCTTCCATCTGAGCGCGACCATGGGAACCCCAAG 61
Db 5 IleProValAlaProTrpProArgVal-----15
QY 62 CCACGNTCTGCGCTGCTGTGCGAGCTGAGCTGGGCGCACTGGAGGGCGTGCC 121
Db 16 ---ArgLeuLysProTrpLeuValAlaGlnValAsnSerCysGlnTyProGlyLeuGln 34
QY 122 TGGGTGAACAAGAGCGGACCGCGTCTCCGATCCCTTGGAGACGAGCGGCTAGCGCAGAT 181
Db 35 TrpValAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly 54
QY 182 GCACAGCAGGAG---GATTTCGGAATCTTCCAGGCTGGCGCGGACCTGCTGCATAT 238
Db 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysLysLysLysTyr 74
QY 239 GTTCCCGGAGGAGGATAAGCAGACCTGCGCAACCTGGAGAGGAATTTCCGCTCTGCCCTC 298
Db 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94

```

```

QY 299 AACCGCAAAAGAGGGTGGCTTATGAGAGGACCGGAGCAGGAC---CCTCAGACCCA 355
Db 95 AsnLysSerArgAspPheArgLeuIleTyArgGlyProArgAspMetProGlnPro 114
QY 356 CATAAATCTACAGATTGTGAACCTCAGGAGTTGGGACATTTCCACGACGACACCTCT 415
Db 115 TyrlsileTyrlGluValCysSerAsnGlyPro-----AlaProThrAspSerGln 131
QY 416 CCG-----CACACCAATGTGGAGGACGAGTACTTCTGATACCCAGGAAGACATCTGGAT 469
Db 132 ProProGluAspTyrlSerPheGlyAlaGlyGluGluGluGluGluGluGln 151
QY 470 GAGTTACTGGGTAACATG-----GTCTTGGCCCCA 499
Db 152 ArgMetLeuProSerLeuSerLeuThrGluAspVallystrpProProThrLeuGlnPro 171
QY 500 CTCCAGATCCGGACCCCGGAGCTGCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTG 559
Db 172 --ProThrLeuGlnProProValValLeuGlyProProAlaProAspProSerProLeu 190
QY 560 CGGAGCCCGAGCTTGGACACATCCCACTCCCTTCCCAACCTGGGGCCCTCTGAG----- 613
Db 191 ---AlaProProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu 208
QY 614 -----AAACCCACTGAACGGCTGTTGGTCCG---GGGGAGAG-----649
Db 209 GluProGlyProLeuProAlaSerLeuProAlaGlyGluGlnLeuLeuProAspLeu 228
QY 650 -----TGGGAGTTTCGAGGTGACAGCTTCTTACCGGGGC 682
Db 229 LeuIleSerProHisMetLeuProLeuThrAspLeuGluLeuLysPheGlnTyArgGly 248
QY 683 CGCCAAAGTCTTCCAGCAGACCATCTCTCCCGGAGGGCGCTGCGCTGTGGGTCGAA 742
Db 249 ArgProProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrlSerGln 268
QY 743 GTG-----GGAGACAGACGCTGCTGGATGGTGGTGGTGGTGGTGGTGGTGGTGG 772
Db 269 LeuGluAlaThrGlnGluGlnValGluLeuPheGlyProIleSerLeuGlu-----Gln 286
QY 773 GTACACTGCCAGACCTGCGCATGCTCCCTGACAGACAGGCGGAGTGTAGTACGTAGG 832
Db 287 ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrlThrAsn 305
QY 833 CATGTGCTGAGCTGCTGCTGGTGGGAGCTGCTCTTGGCGGGCGCGGAGTGGCTCTGG 892
Db 306 GlnLeuLeuAspValLeuAspArgGlyLeuIleLeuGlnLeuGlnAspLeuTyrl 325
QY 893 GCCCAGCGGTGGGCACTGCCACACATCTGG-----GCAGTGAAGCGAGAG 940
Db 326 AlaIleArgLeuCysGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp 345
QY 941 CTGCTCCCAACAGCGGCGCATGGCGCTGATGGCGAGCTCCCAAGGACAGGAAGAGGCG 1000
Db 346 SerCysProAsn-----ProIleGlnArgGluVallystrpLys 358
QY 1001 GTGTTTTCACCTGGGCGCTTCAATTGTAGATCTGTATCTTACCTTCACGGAGAG---AGCGGA 1057
Db 359 LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn 378
QY 1058 CGCTCACACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117
Db 379 ThrProProProGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro 398
QY 1118 TGGACCAAGAGCTGCTGATGCTCAAGGTTCTGCGCCACGCTGCTGCTGCTGCTGCTGCTG 1177
Db 399 ArgGluLysLysLeuIleThrValGlnValProValAlaAlaAlaLeuLeuGlu 418
QY 1178 ATGGCCCGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1237
Db 419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
QY 1238 AGCCA-----CCCACTCTCCCTCAC-----1257

```

```

Db 438 --ProAspLeuLysAspArgMetValGluGlnPheLysGluLeuHisIleTrpGlns 457
QY 1258 -----CTCGACCACTAGCAAGCCCTACCTGAGGACTTGGTGGAGGGCATGGAT 1306
Db 457 erGlnGlnArgGluGlnProValAlaGlnAlaProProGlnGlyAlaGlyLeuGlyValGly- 476
QY 1307 TTCAGGGCCCTGG 1320
Db 477 --GlnGlyProTrp 480

```

RESULT 12

```

IRF5_MOUSE
ID IRF5_MOUSE STANDARD; PRT; 497 AA.
AC PS6477;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN Name=Irfs;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lymph node;
RA Grossman A., Kondo S., Antonio L., Mak T.W.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC - SUBCELLULAR LOCATION: Nuclear.
CC - SIMILARITY: Belongs to the IRF family.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; AF028725; AAB81997.1; -.
DR HSP; P23906; IIRF.
DR TRANSFAC; T05103; -.
DR MGD; MGI:1350924; Irfs.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 16 118 Tryptophan pentad repeat.
FT DOMAIN 9 12 Poly-Pro.
FT DOMAIN 140 148 Poly-Glu.
SQ SEQUENCE 497 AA; 56005 MW; D8BD54DB946E354F CRC64;

```

Alignment Scores:

Align No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score: 9.99e-21	503.50	159	61	189	20
Percent Similarity:	44.35%				
Best Local Similarity:	32.06%				
Query Match:	18.66%				
DB:	1				

US-09-975-253-1 (1-1426) x IRF5_MOUSE (1-497)

```

QY 50 GGAACCCCAAGCCA-----CGNCTCTGCGCTGGTGTGTCGAGCTGGAC 97
Db 7 GlyIleProProProArgValArgValLeuLysProTrpLeuValAlaGlnValAsn 26
QY 98 CTGGGGCAACTGGAGGGCGTGGCTGGTGAACAAGAGCCGCGCTTCGCGATCCCT 157

```

```

Db 27 SerCysGlnTyrProGlyLeuGlnTyrValAsnGlyGluLysLeuPheTyrIlePro 46
QY 158 TGAAGACACGCGCTACCGGAGGATGCACAGAGAG---GATTTCCGAATCTTCCAGGCC 214
Db 47 TrpArgHisAlaThrArgHisGlyProSerGlnAspGlyAspAsnThrIlePheLysAla 66
QY 215 TGGCGCCGAGGCCACTGGTGCATATGTTCCCGGAGGAGTAAGCCAGACCTGCCAACCTGG 274
Db 67 TrpAlaLysGluThrGlyLysTyrThrGluGlyValAspGluAlaAspProAlaLysTrp 86
QY 275 AAGAGGAATTTCCGCTCTGCCCTCAACCGCAAGAGGGTTGGCTTTAGCAGAGACCGG 334
Db 87 LysAlaAsnLeuArgCysAlaLeuAsnLysSerArgAspPheGlnLeuPheTyrAspGly 106
QY 335 AGCAAGGAC---CCTCAGGACCCACATAAAATCTAGAGTTTGTGAACCTCAGAGTTGGG 391
Db 107 ProArgAspMetProGlnProTyrLysIleTyrGluValCysSerAsnGlyProAla 126
QY 392 GACTTT---TCCGAGCCAGACACTCTCCGACACCAATGGTGGAGGCAGTACTTCTGAT 448
Db 127 ProThrGluSerGlnPro-----ThrAspSerTyrValLeuGlyGluGluGlu 143
QY 449 ACCAGGAAGACATTCTGGATGATTACTGGGTAAACATGTTGTGGCCCACTCCAGAT 508
Db 144 GluGluGluGluGluGlnArgMetLeuProGlyLeuSerIleThrGluProAlaLeu 163
QY 509 CCGGAGACCCCA-----AGCTGGCTGTAGCCCTTGAGCCCTGAGCCCTCGCCCTCAGCC 556
Db 164 ProGlyProProAsnAlaProTyrSerLeuProLysGluAspThrLysTrpProAla 183
QY 557 CTGCGAGAGCC---AGCTTGACAAATCCCACTCCCTCCCAACCTG-----GGGCC 607
Db 184 LeuGlnProProValGlyLeuGlyProProValProAspProAsnLeuAlaPro 203
QY 608 TGTGAGAACCCA-----CTGAAGCGGCTGTTGGTCCGCGGG-----643
Db 204 SerGlyAsnProAlaGlyPheArgGlnLeuLeuProGluValLeuGluProGlyProLeu 223
QY 643 -----643
Db 224 AlaSerSerGlnProProThrGluProLeuLeuProAspLeuLeuIleSerProHisMet 243
QY 644 ---GAAGAGTGGAGTTCGAGTGACAGACCTTCTACCGGCGCCGCAAGTCTTCAGAG 700
Db 244 LeuProLeuThrAspLeuGluLeuIleLysPheGlnTyrArgGlyArgAlaProArgThrLeu 263
QY 701 ACCATCTCTCCCGAGGCGCTCGGCTGGTGGGTCGCAAGTG-----745
Db 264 ThrIleSerAsnProGlnGlyCysArgLeuPheTyrSerGlnLeuGluAlaThrGlnGlu 283
QY 746 -----GGAGACAGACCGCTGCTGGATGGCCAGTGCACATCCAGACCCCT 790
Db 284 GlnValGluLeuPheGlyProValThrLeuGlu-----GlnValArgPheProSerPro 301
QY 791 GGCATGTCCTACAGACAGGAGTGCATGAGCTACGTGAGGCATGCTGCTGAGCTGCTG 850
Db 302 GluAspIleProSerAspLys---GlnArgPheTyrThrAsnGlnLeuLeuAspValLeu 320
QY 851 GGTGGGGAGTGGCTCTCTGCGCGCGCGGCGGCGAGTGGCTGTCGGGCCAGCGCTGGGGCAC 910
Db 321 AspArgGlyLeuIleLeuGlnLeuGlnAspLeuTyrAlaIleArgLeuCysGln 340
QY 911 TGCACACATACTGG-----GCAGTACGAGGAGGCTGCTCCCAACACAGCGGG 958
Db 341 CysLysValPheTrpSerGlyProCysAlaLeuAlaHisGlySerCysProAsn-----358
QY 959 CATGGGCTGATGGCGAGGTCCTCCCAAGGACAAAGAGGAGCGCTGTTGACCTGGGGCCC 1018
Db 359 -----ProIleGlnArgGluValLysThrLysLeuPheSerLeuGln 373
QY 1019 TTCATTGTATCTGATTACCTTTCAGGGAAGGA---AGCGGAGCGCTCACCACCTATGCC 1075

```

Db 374 PheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsnTrpProProPheGlu 393

Qy 1076 CTCTGGTTCTGTGGGGAGTCA TGCCGCCACGACCAGCGCTGGACCAAGAGGCTCGTG 1135

Db 394 IlePhePheCysPheGlyGluGluTrpProAspValLysProArgGluLysLysLeuIle 413

Qy 1136 ATGTCGAAGTTGTGCCACCTGC TCAGCGCCTTGGTAGAAATGGCCCGGTAGGGGGT 1195

Db 414 ThrValGlnValProValAlaAalargLeuLeuLeuGluMetPheSer--GlyGlu 432

Qy 1196 GCCTCCTCCCTGGAGAATACTGT GCGACTGCACATTTC AACGCCACCATCTCCCTC 1255

Db 433 LeuSerTrpSerAlaAaspSerIleA argLeuGlnIleSerAsnProAspLeuLysAspHis 452

Qy 1256 ACCTCCGCCACGATCAAG----- GGCTACTCTCAGGAC 1288

Db 453 MetValGluGlnPheLysGluLeuH isLeuTrpGlnSerGlnGlnLeuGlnPro 472

Qy 1289 TTGGTGGAG-----GGCATGGAT ---TTCCAGGGCCCT 1318

Db 473 MetValGlnAlaProProValAlaGly leuaspAlaSerGlnGlyPro 498

```

RESULT 13
IRF5 HUMAN
ID IRF5 HUMAN STANDARD; PRT; 498 AA.
AC Q13568; Q9BQF0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN Name:IRF5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP Grossman A., Mitrucker H.W., Lantonio L., Mak T.W.;
RA Submitted (APR-1996) to the ENBL/GenBank/DBJ databases.
RL
[2]
RN SEQUENCE FROM N.A.
RP

```

TISSUR=Kidney; PubMed=12477932; DOI=10.1073/pnas.2426039899;

MEDLINE=22388257; Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L., Datchenko L., Maruina K., Farmer A.A., Casavant T.L., Schetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Mullany S.J., Raba S.S., Locuelli N.A., Peters G.J., Malek J.A., Gunaratne P.H., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Hulyk S.W., Richards S., Worley K.C., Hale S., Garcia A.J., Lu X., Gibbs R.A., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.; Scherch A., Schein J.E., Jones S.J.M., Marra M.A.; "generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

-!- SUBCELLULAR LOCATION: Nuclear.

-!- SIMILARITY: Belongs to the IRF family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

or send an email to license@isb-sib.ch.

```

-----
CC      EMBL; U51127; AAA96056.1; -.
DR      EMBL; BC004201; AAH04201.1; -.
DR      EMBL; BC004139; AAH04139.1; -.
DR      PIR; G02474; G02474.
DR      HSSP; P23906; 1IRF.
DR      TRANSFAC; T05104; -.
DR      Genew; HGNC:6120; IRF5.
DR      MIM; 607218; -.
DR      InterPro; IPR001346; IRF.
DR      InterPro; IPR008984; SMAD_FHA.
DR      InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR      Pfam; PF00605; IRF; 1.
DR      PRINTS; PR00267; INTERREGFCT.
DR      ProDom; PD002355; IRF; 1.
DR      SMART; SM00348; IRF; 1.
DR      PROSITE; PS00601; IRF; 1.
DR      DNA-binding; Nuclear protein; Transcription regulation.
KW      DNA BINDING 16 118      Tryptophan pentad repeat.
FT      DOMAIN 142 149      Poly-Glu.
FT      CONFLICT 161 175      EDVKWPTTQPPTLR -- DAVQSGPHMTFYSLLKEDVKW
                                     (in Ref. 1).
FT      SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

Alignment Scores:
Pred. No.: 1.69e-20
Score: 499.50
Percent Similarity: 43.60%
Best local Similarity: 31.59%
Query Match: 18.51%
DB: 1

```

US-09-975-253-1 (1-1426) x IRF5_HUMAN (1-498)

QY	2	GTTCAGCTGTCGGCAGCCGCCGACCTTCATCGTAGCCGGACCATGGGAACCCCAAG	61
Db	5	tleProValAlaProThrProArgVal	15
QY	62	CCACGGNTCCCTGCCTGGCTGGTGTGCGAGCTGGACCTGGGGCAACTGGAGGGGTGGCC	121
Db	16	---ArgLeuLeuProThrIleuValalaglnValAsnSerCysGlnTyrProGluLeuGln	34
QY	122	TGGGTGAACACAGAGCCGACGCGCTTCGCGATCCCTTGGAGACACGCGCTACGCGCAGAT	181
Db	35	TrpValAsnGlyGluLysLeuPheCysIleProThrArgHISAlaThrArgHisGly	54
QY	182	GCACAGCAGGAG---GATTCGGATCTTCCAGGCTGGCGCGAGCCACTGGTGTCATAT	238
Db	55	ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTyr	74
QY	239	GTTCCTGGGAGGATGAAGCCAGCACCTGCGCAACTCGAAGAGAGAATTCCTGGCTCTGGCTC	298
Db	75	ThrGluGlyValaspGluAlaaspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu	94
QY	299	AACCGCAAGAAGGGTGGCTTTAGCAGAGGACCGGACCAAGGAC---CCTCACGACCCA	355
Db	95	AsnLysSerArgAspPheArgLeuIleTyrAspGlyProArgaspMetProGlnPro	114
QY	356	CATPAAATCTACAGTTTGTACTCAGAGTTGGGGACTTTTCCCGCCGACGACACCTCT	415
Db	115	TyrLysIleTyrGluValCysSerAsnGlyPro-----AlaProThrAspSerGln	131
QY	416	CCG-----GACACCAATGGTGAGGACGACTCTGTGATACCCAGGACACATTCTGGAT	469
Db	132	ProGluaspTyrSerPheGlyAlaGlyGluGluGluGluGluGluGluGln	151
QY	470	GAGTTACTGGGTACATG-----	487
Db	152	ArgMetLeuProSerLeuSerLeuThrGluAspValLysTrpProThrLeuGlnPro	171
QY	488	-----GTGTTGGCCCCACTCCCGAGATCCGGGACCCCAACGCTGGCT	529
Db	172	ProThrLeuArgProProThrLeuGlnPro---ProThrLeuGlnProProValValLeu	190

Db 426 SerGluMetPheSerClyAspCysThrArgSerPheAspSerGlySerValArgLeuGln 445
 QY 1229 ATTTCACAGCAGCACCATCTCCCTCCTCAGCAGCAGTACACAGGCTACCTGACGAG 1288
 Db 446 lIeSer-----lIleProAspIleLysAspAsnIleValThrHisLeuLysGln 461
 QY 1289 TTGGTCGAGGCATGGAT---TTCCAGGGGCGCTGGGAGAGAGCTAGCGCTCGCTCT 1342
 Db 462 LeuTyArgLeuLeuGlnThrHisGlnGlyPro---GluAsnTrpProLeuProPro 479

RESULT 15
 AAH56772 PRELIMINARY; PKT; 492 AA.

ID AAH56772; AC AAH56772; DT 24-MAY-2004 (TrEMBLrel. 27, Created)
 DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 24-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein zgc:63500.
 OS ZGC:63500.
 GN Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN [1]
 RX SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Kaba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RX SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RA Strausberg R.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL; BC056772; AAH56772.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 492 AA; 54876 MW; 98EF8FD3FE1F589B CRC64;

Alignment Scores:
 Pred. No.: 2 86e-18 Length: 492
 Score: 460.50 Matches: 146
 Percent Similarity: 42.69% Conservative: 67
 Best Local Similarity: 29.26% Mismatches: 185
 Query Match: 17.06% Indels: 101
 DB: 2 Gaps: 21

US-09-975-253-1 (1-1426) x AAH56772 (1-492)

QY 65 CGGNTCTCGCCCTGGCTGGTGTCGACCTGGAGCTGGGGCACTAGGAGCGGTGGCGCTGG 12
 Db 9 ArgLeuLysProTrpLeuValSerGlnValAspAsnAlaThrPheProGlyLeuValTrp 28
 QY 125 GTGAACAGAGCCCGACCGCGCTTCGCGATCCCTTTGGAAGCAGCGGCTACGCGAGGATGCA 18

This Page Blank (uspto)